DOCUMENT RESUME

ED 288 763 SO 018 390

AUTHOR Wunderlich, Gene, Ed.
TITLE Ideas, Facts, and Choices: Three Lectures on Philosophy and Applied Economics.
INSTITUTION Economic Research Service (DOA), Washington, D.C.
REPORT NO AGES870512
PUB DATE Jun 87
NOTE 53p.
PUB TYPE Viewpoints (120) -- Collected Works - General (020)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Economic Change; *Economic Factors; *Economics; Ethics; Logic; *Philosophy
IDENTIFIERS *Economic Theory; *Economic Trends

ABSTRACT

Four philosophers, familiar with the work of economists, comment on the foundations and substance of applied economics. Joseph Pitt, in "Philosophy in Applied Economics," focuses on rationality and rational choice as a foundation for economics and stresses the idea of relevance in relating data to applied economics. He believes rational decision making is based on the notions that choice is necessary and possible and that behavior is predictable. However, Pitt considers the economic model a poor predictor of behavior. Allen Buchanan, in "Comparative Efficiency Judgements, Economics, and Political Philosophy," develops a rationale to show that there are fundamental obstacles to efficiency comparisons of entire economic systems and reviews market suppositions developed by economists and ethicists. Douglas MacLean and Claudia Mills examine the intellectual hazards of confusing normative and empirical statements about behavior in "Normative and Empirical Issues in Economic Theory: A Philosophical Examination." They consider prospect theory a possible solution to problems inherent in expected utility theory. All three papers provide comments on the model-making and testing processes of applied economics. "Postscript: The New Spirit of Inquiry in Economics and Philosophy," by Michael McPherson, considers the effects of economists' and philosophers' collaborative efforts. (JHP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED288763

United States
Department of
Agriculture

Economic
Researc
Service

**Natural
Resource
Economics
Division**

Ideas, Facts, and Choices

Three Lectures on Philosophy and Applied Economics

SO018390

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

IDEAS, FACTS, AND CHOICES: THREE LECTURES ON PHILOSOPHY AND APPLIED ECONOMICS. Gene Wunderlich, editor. Natural Resource Economics Division, Economic Research Service, U.S. Department of Agriculture. Staff Report AGES870512.

ABSTRACT

Four philosophers comment on foundations and substance of applied economics. Pitt focuses on rational decision making and stresses the importance of choice although he faults the economics model as a predictor of behavior. Buchanan argues that comparisons of whole economic systems on the basis of efficiency are doubtful if not impossible. MacLean and Mills examine the hazards of confusing normative and empirical statements about behavior. All papers provide constructive comments on the model making and testing processes of applied economics.

****
* This report was prepared for limited distribution to the research *
* community outside the U.S. Department of Agriculture. *
****

1301 New York Avenue, NW.
Washington, DC 20005-4788

June 1987

CONTENTS

Page

INTRODUCTION
by Gene Wunderlich 1

PHILOSOPHY IN APPLIED ECONOMICS
by Joseph C. Pitt. 4

COMPARATIVE EFFICIENCY JUDGMENTS, ECONOMICS, AND POLITICAL PHILOSOPHY
by Allen Buchanan 19

NORMATIVE AND EMPIRICAL ISSUES IN ECONOMIC THEORY: A PHILOSOPHICAL EXAMINATION
by Douglas MacLean and Claudia Mills 33

POSTSCRIPT--THE NEW SPIRIT OF INQUIRY IN ECONOMICS AND PHILOSOPHY
by Michael McPherson 48

Ideas, Facts, and Choices

Three Lectures on Philosophy and Applied Economics

INTRODUCTION

by Gene Wunderlich

> As we all know, a little knowledge is a dangerous thing. But a great deal of highly specialized learning is also a dangerous thing and may be even more dangerous than a little learning. . .
> (Aldous Huxley, The Human Situation, 1977)

Agricultural economists, by the nature of their work, tend to be specialists. Indeed, much of their comparative advantage, and often their rewards result from specialization. However, even as specialists, they find it valuable occasionally to place their processes and products in a larger environment. The observations of others with different models and assumptions about the world, expressed in different vocabularies, can be useful in creating that larger environment.

Three philosophers familiar with the work of economists were invited to discuss their ideas on applied economics with the staff of the Economic Research Service in the spring and early summer of 1986. The papers are arranged in the order of their presentation as lectures, and run broadly from the general to the specific.

Joseph Pitt, in describing what philosophers do, provides a setting for his discussion of rational choice and for the papers by Buchanan and MacLean/Mills. He draws on Wilfrid Sellars' characterization to encapsulate the aim of philosophy as a way of "knowing how things hang together." Pitt blends several areas of inquiry with philosophy, which he says "does not appear essentially different from any other discipline. We all try to clarify the meaning of what we are saying and to justify the truth or falsity of our claims." Out of this general setting, he focuses on rationality and rational choice as a foundation feature of economics. But he does not confine rationality to ratiocination. He embodies notions of relevance in relating data to applied economics. Truth of data is "contextualized" by theory: that is, data are not true in some absolute sense but in relation to a theory and presumably in relation to an environment of observation.

Gene Wunderlich is an agricultural economist with the Economic Research Service, U.S. Department of Agriculture.

Rational decision making, according to Pitt, is based on a set of assumptions. First, rational decision making presupposes that choice is necessary, not only possible. Second, rational choice is possible. It is this second assumption that gives Pitt some problem with the basic model of choice in economics and, in a sense, joins Arrow and others on the issue of ordered preferences. Are ordered, consistent preferences necessary for the success of economic models--particularly in their empirical representation or test? With a third assumption, predictability of behavior, Pitt seems to criticize the economic model. But must a model predict perfectly to meet a standard of rationality? Perhaps the issue of rationality depends on definitions of determinate and probabilistic prediction. The reader will have to decide how conclusive Pitt's arguments on rationality are. Clearly, however, the idea of choice and its empirical representation is at the heart of economics as a social or behavioral science.

From a more comprehensive treatment in his book, Ethics, Efficiency and the Market, Allen Buchanan develops an argument to show that there are fundamental obstacles to comparisons of entire economic systems on the grounds of efficiency. He reviews the essential features of, and difficulties with, Pareto standards in production and distribution. The coexistence of states of production with states of distribution tends to render Pareto optimality inconclusive or trivial. From his review of the Pareto criteria, Buchanan finds the efficiency comparisons no less subject to relativism than comparisons by ethical criteria. He concludes that intersystemic comparisons of economies on efficiency standards are not useful, especially where the systems to be compared are radically different.

In making his argument, Buchanan reviews the important suppositions about the market made by economists and by ethicists. He emphasizes the "interplay between efficiency assessments and moral judgments." Although he concentrates most attention on the familiar limitations of the Pareto criteria, he supports the idea that reason can be applied to ethical issues. He rejects the ethical relativism of some economists, that is, that there is no right or wrong, only revealed preferences. That communities and societies are able to exist and even develop attests to the existence of shared goals and modes of behavior. Thus it may be impossible to make direct interpersonal utility comparisons, but it is not impossible to make ethical judgments about behavior, including distributions of wealth and income.

Douglas MacLean, with coauthor Claudia Mills, examines the intellectual hazards of confusing normative and empirical statements about behavior. While they attribute such confusions to philosophers and scientists generally, their case is built on issues in economic policy. They concentrate on developments in decision theory and indicate the importance of distinguishing theories of ideal behavior from facts about actual behavior. They recount the evolution of decision theory from a set of intuitively appealing axioms explaining attempts to maximize utility. As conflicts arose between empirical facts and the "as if" theory of maximizing expected utility, economists tried to reconcile facts and theory.

After their effort to clarify the distinction between normative and empirical claims, MacLean and Mills briefly review prospect theory as a response to problems inherent in expected utility theory. Behavior is explained not by some defined utility function but by gains and losses from some perceived position. Gains and losses are asymmetrical, so solutions are determined by the way problems are framed. From their excursion into decision theory, the authors draw upon the need to separate normative and behavioral theories. They

illustrate in terms of some traditional economic cases of "buying in" or "buying out," that is, compensation versus acquisition. Finally they address the framing of entitlements or rights as moral issues. Is the moral right to hold a property of greater value than the property itself? Is the right to hold an object equivalent to the right to acquire it? As did the preceding papers, the one by MacLean and Mills stimulates more questions than it answers.

All of the authors cheerfully complained that the scope of their assignment, philosophy of applied economics, was too broad and they had narrowed their responses to manageable topics: Pitt to rational choice, Buchanan to intersystemic comparisons of economies, and MacLean/Mills to behavioral decision theory. From a synthesis of their topics and paper contents we arrived at the title Ideas, Facts, and Choices--a bit pretentious perhaps, but not if understood in the context of working notions, not final words.

Travel is said to be broadening. The lectures were intended as an excursion into the territory of philosophers and thus to broaden our perspectives of, and ideas about, economics. It is in this spirit of travel and exchange of ideas that the three lecturers agreed to put their remarks in writing. Michael McPherson, coeditor of "Economics and Philosophy" and a resident of Washington, DC, at the time of the lectures, provides an insightful postscript. Whether you agree or disagree, we hope you will find these papers useful.

PHILOSOPHY IN APPLIED ECONOMICS

by Joseph C. Pitt

The large topic of philosophy in applied economics may be made somewhat more manageable by concentrating on the problem of rational choice; more specifically, the philosophical presuppositions underlying rational choice. But first, two preliminary issues should be considered. These are: (1) the nature of philosophy and its relevance to economics, and (2) the difference between theoretical and applied economics. First, the philosophy, then the more specific questions about economics and rational choice.

A Very Brief Introduction To Philosophy

As is well known, the term "philosophy" comes from two Greek words, "philo" and "sophia." The meaning of "philosophy," therefore, is "love of wisdom." But the etymology of the word does not say much about philosophy's domain of inquiry or its aims and concerns. So let us consider a second attempt at characterizing philosophy. Here is Wilfrid Sellars'.

> The aim of philosophy, abstractly formulated, is to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term. Under 'things in the broadest possible sense' I include such radically different items as not only 'cabbages and kings', but numbers and duties, possibilities and finger snaps, aesthetic experience and death. To achieve success in philosophy would be, to use a contemporary turn of phrase, to 'know one's way around' with respect to all these things, not in that unreflective way in which the centipede of the story knew its way around before it faced the question, 'how do I walk?' but in that reflective way which means that no intellectual holds are barred. (9) 1/

On the surface, Sellars' account may not appear to advance us much beyond the lover of wisdom routine, but it does. Sellars provides us with a couple of important signposts in our journey toward figuring out philosophy. The first is his last point: no intellectual holds barred. The second is his condition of success: to know one's way around. To say no intellectual holds are barred means all knowledge claims are open to the most intense scrutiny with respect to their content and their justification. Philosophers are infamous for asking such questions as "what does that mean?" and "what is your justification for holding that view or making that claim?" The aim of such questioning is to

---

Joseph C. Pitt is a Professor of philosophy at Virginia Polytechnic Institute and State University.

1/ Underscored numbers in parentheses refer to sources cited in References at the end of this article.

clarify what is being said and its basis in fact or reason. Beyond clarifying the meaning of individual claims, there is also the problem of understanding these claims in the context of what else is being proposed or defended or maintained. The notion of "knowing one's way around" entails understanding individual claims in their appropriate context.

Philosophy, as described above, does not appear essentially different from any other discipline. We all try to clarify the meaning of what we are saying and to justify the truth or falsity of our claims. Therefore, it should seem to be the case that everyone qualifies as a philosopher. However, Sellars spoke initially of the aim of philosophy as understanding of how things in the broadest possible sense hang together. "In the broadest sense" should be construed in the broadest sense. In other words, if Sellars is right, philosophy is concerned with life, the universe and everything, and its goal is to make sense of all of it, together. That is a bit more to handle.

Mature And Speculative Philosophy

The problem of trying to make it all make sense is complicated by the indirect way philosophers deal with the world. The subject matter of mature philosophy is the totality of human effort to make sense of the world. Mature philosophy does not deal with the world per se, but rather with the efforts of others to deal with the world. Mature philosophy is, in this very restricted sense, parasitic on the results of other disciplines. Charitably stated, mature philosophy is a second-order enterprise, whose domain is not the world, but the confusions that arise from our attempts to understand the world.

There is, in addition to mature philosophy, a different kind of enterprise called speculative philosophy that attempts to explain the world directly. Despite its coming last in this discussion, speculative philosophy is actually first in the history of philosophy. Earlier I called mature philosophy parasitic in a very restricted sense. The reason for the restriction is that it simply cannot be true if understood from the point of view of history of thought. Beginning with the Greeks of ancient Athens, all but one of the major fields of human inquiry derive from the results of prior speculative philosophical investigation. Mature philosophy has as its domain the results of that speculation as currently represented in our modern disciplines. Furthermore, the process of speculation leading to an intellectual discipline continues today. One of the more recent examples is the birth of linguistics out of the concerns of ordinary language analysis. The one area that is not so directly related to speculative philosophy is religion. While I do not intend to argue the case here, I view philosophy as rational man's response to religion. It is an attempt to come up with an account of the world and our place in it based on reason and the rejection of mystery.

The manner in which other fields derive from philosophy is worth a paper in itself. But, here a summary must suffice. Before a science is possible, certain preliminary questions must be clarified. That is, before empirical research can take place, preliminary work is needed to isolate the empirical dimension of the issues. Philosophers, in their early speculative attempts to explain how things hang together, often propose what I will call proto-explanations. These are prescientific but nonreligious accounts of why things hang together in the way they appear to. These usually first appear rather grand and eloquent, but under scrutiny come to be seen as vague and ultimately more confusing than helpful.

As an example of grand speculative proto-explanations, consider Plato's account of the world as a shadow of Reality. On the surface, it manages an interesting attempt at explaining the world and our knowledge of it. However, under examination it suffers because of a number of problems that can only be solved at the expense of the initial insight that gave the system its power. Nevertheless, Plato's proto-explanations still warrant examination. He articulated a point we now take for granted: what we see is not necessarily all that is there.

Plato also gave form to one of the more enduring questions in philosophy: what is the status of mathematical entities? Do they exist for real, or are they only a product of man's creative power? If they exist, in what manner do they exist? Surely, numbers are not like tables and chairs. This problem takes on some urgency today as we see increasing quantification in all the sciences. Our fascination with numbers often tends to obscure the question of what they mean. In both physics and economics, we appear to be pushed to the conclusion that ultimate reality is a set of mathematical relations. For materialists like myself who like their creature comforts, there is something disturbing in the notion that all is number.

One of the first jobs of a philosopher is to take the vaguely worded proto-explanations we have offered and refine both the questions and the sense of what would be a relevant answer. In large part, this amounts to sorting out what constitutes good reasons for an answer to a specific question. This, in turn, requires finding out what the question is really asking. When this has been accomplished we have the basis for pointing to or isolating a domain of inquiry. When the nature of the domain is secured and the kinds of questions appropriate to it finally determined, we have the possibility of empirical research. While philosophers do not usually engage in empirical research, the kind of prior conceptual analysis they perform makes it possible in many cases for productive empirical research to emerge. Thus, we find metaphysics leading to physics, philosophy of language producing linguistics, and questions in ethics and political philosophy leading to political economy and finally economics itself. In that sense then, all of science is parasitic on philosophy. The nice way to say this is that philosophy is the queen of the sciences.

Philosophy of Science and Rationality

But this is not where the story ends. We start with speculative philosophy and refine its insights, thereby making it possible for an empirical domain of inquiry to emerge. But we end with mature philosophy. Mature philosophy, as I see it, is primarily concerned with assessing the methodological dimensions of the sciences to which philosophy gave birth. In particular, philosophy of science is concerned with the problem of justification in the empirical sciences. This is not to say that the business of philosophy is to tell scientists how to do science. It is rather to determine when we should believe what scientists tell us. To make such a determination requires that we have some understanding of rationality.

Philosophy is the rational person's response to the unintelligible. As such, part of the philosophical enterprise, if not its heart, is the continuing attempt to explain what such a response ought to be, that is, what it means to be rational. Regardless of the specific approach one takes to the problem of defining "rationality," it is agreed that to be rational is to be in a position

to provide good reasons for one's actions, proposals, and so on. The philosopher's questions and constant demands for reasons and justification presuppose, furthermore, a standard, a criterion of adequacy for good reasons. Thus, the philosophical quest is for standards of good reasons. The philosopher begins his quest in good faith, with the assumption that to make sense of the question we need to find the criterion by which to judge an answer adequate.

First, there is the question of where to begin. Should philosophers look to the kinds of answers economists and other scientists have provided and attempt to elicit a standard of good reasons from those pronouncements for each of their separate domains? Or should they formulate one abstract standard for the whole set of domains which purport to be of one kind, for example, the sciences? In other words, a rather basic preliminary decision needs to be made here: start from the bottom, with the concerns of the scientists, or start at the top with some general characterization of science and work down to science?

If we start at the bottom and consider the kinds of reasons economists give and accept as legitimate, it does not follow that we have an answer to the problem of what the appropriate standard ought to be. For as Hume reminded us, what is the case does not entail what ought to be the case. The fact that economists reason in a certain way is an interesting fact about economists. "But," the philosopher asks, "are the economists's reasons well grounded?" In other words, the question concerning the status of good reasons remains regardless of what people do. It is a question about what ought to be good reasons, it is not a question of fact. Thus, if we start at the bottom, we still have the question of what constitutes an appropriate standard for good reasons in economics.

We could, on the other hand, start at the top. This would involve setting out a definition of "science" and then elaborating the standards under which reasoning in a given domain qualifies as "scientific." In so doing we would effectively beg an important issue, namely: are all sciences to be judged by the same standard? Being scientific does not necessarily mean using just one criterion of good reasons for all sciences. To insist on such a universal standard would, among other things, restrict scientific inquiry to accepting only one form of evidence for all sciences. Furthermore, a general or universal definition of "science" may not embrace all those activities we consider relevant. What if a definition of science looks good but, for example, won't accommodate sociology or genetics? If they do not fit the definition, it does not follow they should not count as sciences. Rather, so much the worse for the definition.

If a definition of science fails to do justice to what we think ought to count as sciences, our search for good reasons is in jeopardy. There is a general solution to problems of this type: if the definition does not work in the face of obvious counterexamples, change it. Much of what goes on in philosophical circles is a form of trial and error, readjust-the-definition-in-the-face-of-failure. But the problem of adequately defining science faces serious problems which cannot be solved in that fashion. There are good reasons to think not only that no single account of science currently adhered to adequately captures all of what we accept as science, but that no such single account could. There are many aspects of the situation which contribute to this conclusion.

For one thing, as a close examination of the history of science shows, science is an ongoing process in which both domains and methods change over time in

light of new theories and evidence. To assume that such a dynamic process can be captured by one universal definition seems unreasonable.

Second, there is considerable debate within the scientific community over the relationship among the sciences. Consider, for instance, the longstanding argument between biologists and physical scientists over whether biology can or should be reduced to physics. In part, the solution to this dispute turns on the acceptability of a uniform standard of explanation for both sciences. A uniform standard of explanation means a uniform account of good reasons. But that is where we came in.

Third, standards of adequacy are often subjects for debate within specific sciences. The question of whether biology can or should be reduced to physics is, for example, not merely a worry of philosophers, it is also an issue within the biological community. In addition, there is in biology some serious question over the status of genetics. Is it fundamental in biology? Economics has similar problems. For example, where does the work of the new experimentalists fit? What is the relation between pure and applied economics?

The social sciences face still another problem. Are they to employ the same standard of good reasons as the physical sciences? And within the social sciences, there is some discussion as to whether they can all be held to the same standard, be it internal to the social sciences or external with respect to the physical sciences. Put cleanly: is physics an appropriate model for economics?

So far we have looked at some of the problems of establishing what constitutes good reasons when we start at the top with an abstract definition of "science," and when we start at the bottom with the informal beliefs of practicing scientists. There is yet another way to begin the search for standards of good reasons. The philosopher can simply attempt to reason out what ought to be an appropriate standard for economics, given some abstract characterization of what economics is about. But surely, by now one should be able to anticipate the kinds of objections that will be raised against this move. To begin with, why should a philosopher do this and not an economist? Is there only one characterization of economics? Will the goal of economics be characterized differently depending on one's ideology? How realistic is a philosopher's understanding of economic reality? As we shall see in the following section, despite its dangers, there is something to be gained from taking this abstract approach to the search for good reasons in economics, provided one remains at a level of discourse reasonably removed from the details of particular economic theories.

We have arrived, therefore, at the point where we should leave our brief hob-nailed forced march through the garden of general philosophy and turn our attention to some specific questions in the philosophy of economics. In summary, philosophy has a long history and comes in various guises. Its aim is to know one's way around, meaning by that, to have a robust account of good reasons why things hang together.

Pure and Applied Economics

Because our general topic is philosophy in applied economics, it seems appropriate to pay some attention to the distinction between pure and applied economics. There are many ways to explain the difference. A simple approach

might focus on the role of variables. In pure economics, we do not have to pay attention to the variables. That is, in pure theory the objective is to define the problem area, identify the variables, and sketch a series of strategies for arriving at solutions. In applied economics, we have to pay more attention to the variables, by deciding, for instance, how to weight those variables. This means one problem is to find good reasons for choosing to weight the variables in the manner we do. We will discuss some issues associated with this aspect of applied economics later. But first we need to examine a still more fundamental set of problems.

We have to pay attention to the variables in applied economics because ultimately we are concerned with policy formation. The strength of policy recommendations, and the probability of their success if implemented, rests not only on the reasons we advance for weighting our variables, but on the manner in which we employ our data and the reliability of the procedures we use to generate it. In other words, the reliability of the data is the basis for sound policy. That presents us with two questions: (1) Can we determine the accuracy of our data? and (2) How do we know we have the appropriate or relevant data?

Question 1 can be restated as the problem of truth. How do we guarantee that the data we gather and generate are true? There are two answers to this question, one short and one long. The short answer is: we cannot. The long answer comes to the same conclusion, but takes us all the way around Jack Robinson's barn. It is, nevertheless, worthwhile to sketch the relevant moves in the argument so as to give the flavor of the problem.

The main reason we cannot be certain of the truth of our data is that truth is always contextualized, that is, relative to a theory. A theory is a set of sentences which embodies assumptions about the population and organization of a domain of inquiry. If you change the assumptions you change the kinds of things one looks for, as well as what one can say about them.

The role of theory is fundamental in science. Science proceeds through the articulation and refinement of theories. Science is not a neutral and dispassionate search for truth among the lilies, it is a deliberate attempt to determine the merits of following a certain set of assumptions. In short, there is no such thing as what used to be called the five-step scientific method, which began with the injunction "make observations." Scientists are not pure observers. Those who speak of the scientific method leave the impression that scientists approach the world free of prejudice, innocent and pure. That simply is not the case. Consider only the fact that a considerable part of the education of a scientist consists of learning the relevant theories in the field and becoming acquainted with the techniques for applying them. When the scientist begins to make observations, it is the context of attempting to prove some hypothesis which has been formulated within the context of a theory. Scientists simply do not go out into the field and make observations of anything, everywhere, with no restraints. They are restrained by the assumptions of a theory and it is in this sense that data are always relative to a theory.

Once we realize data are contextualized in this way, it becomes easier to understand why their truth is so difficult to determine. The determination of the truth of a sentence is an attempt to check for the correctness of fit between what the sentence claims and the way the world is. To accomplish this requires stepping outside the theory in which the sentence is formulated and

comparing it to the world. In other words, there must be a framework which is neutral with respect to the two things being compared for the comparison to take place. This framework will, in turn, be another theory. Furthermore, this broader theory must contain certain assumptions about what an adequate mapping of a sentence onto the world must look like, if it is going to be useful for determining the truth of a claim. This assumption is a theoretical assumption. There is no fact of the matter here. That is, if we assert that a sentence is true if it correctly describes the world, and it describes the world in subject-predicate terms, we thereby endorse a rather specific and dubious assumption about the world. This assumption amounts to the claim that the world is composed of objects and their properties. Thus, on this view, a correct sentence says of the objects in the world that they have some specific property, for example, the cat is black.

But it is not clear this is the proper way to parcel the world, that is, into objects and properties. It could be that this assumption merely reflects a feature of our peculiar physiological makeup. Surely we do not want to claim our physiological constitution has some kind of metaphysical priority. That would be to exceed even the limits of hubris. The world could have other structures. In fact, modern theoretical physics even suggests the universe is a complex of interacting forces, rather than objects. Thus, the correct view might be a universe of perpetual process, rather than one of discrete episodes and facts.

It looks as if a truly neutral framework is impossible to find if each framework embodies assumptions which cannot be tested except by appeal to yet another framework which in turn embodies assumptions which cannot be tested, and so on. It, therefore, follows that such a neutral framework is impossible. We are unable to determine the truth of any given sentence except in this most highly contextualized fashion.

The problems of determining the relevance of data to a theory are similar to the problems facing the determination of truth. The problem of relevance is a larger version of the problem of confirmation. The latter concerns the question of how one determines whether data support or fail to support a theory. Prior to that, however, one must be able to show that the data are relevant. The question of relevance requires a set of standards, independent of the theory, which can be invoked for the purpose of showing how the data bear on the theory. To do this requires that the standard be formulated in a broader context than the theory alone for the comparison between theory and data to take place. The question that puts the whole process in doubt centers on the appropriateness of the standard invoked to determine relevance. To evaluate the standard we must have recourse to a broader framework, and so forth.

In sum, there are severe problems facing the efforts to show our data are relevant and true. But until we solve such problems, the reliability of our policy recommendations is suspect. At present, it appears the best we can do is engage in a continuing process of readjusting our assumptions in the light of whatever results we obtain. This process will not produce truth in any determinate sense. Rather, it will produce at best a form of coherence among our assumptions and our beliefs. This coherence may, in fact, serve us better than any claim of absolute truth in the long run when it comes to policy formation, for many problems of policy formation seem to derive from the conflict among assumptions. If we can at least sort out our assumptions, we might be in a better position to argue for the rationality of our decisions.

The Economic Idea of Rational Decisionmaking

The idea of rational choice in the context of economic theory appears in at least three discrete areas: (1) the concept of economic man, (2) the construction of theory, and (3) the formulation of policy recommendations. Drs. Buchanan and MacLean address numbers two and three in their papers. This paper is restricted to the philosophical presuppositions underlying rational choice as they are exhibited in the concept of economic man, or preferably, homo economicus.

The three points of concern to us were stated in the above order deliberately. Recall that these were the concepts of homo economicus, theory development, and policy formation. As presented they are progressively dependent. That is, working backwards now, the formulation of policy recommendations takes place in a context provided by a theory. The theory provides the parameters we manipulate when formulating our recommendations. As policymakers we actually use a number of theories. For example, if our policy objective is to increase personal savings, we employ theories about consumer behavior, the market, and the role of government. These theories give us the variables we need to manipulate. Our objective may or may not be theory generated. But the objective will be the solution to a problem. To say that the objective may not be theory-generated is merely to recognize that while the problem is formulated in the context of a theory, the need to solve that problem, as so described in the language of the theory, may be generated by some nontheoretical circumstance.

In what follows, some methodological issues will be developed with simply stated examples and scenarios. The issues exposed here reveal some philosophical assumptions of decisionmaking without concern for details of particular models or the merits of some variety of theory. These assumptions about decisionmaking are antecedent to the problems of discrete models.

Assume you find out by a series of individual encounters and reports of others that lots of people are out of work. You determine that the initial problem is to find them employment. If that is the problem, then the first step in solving it is to find out why there are no jobs for these people who used to have jobs. We can say that the discovery of people out of work does not indicate a theory-generated problem. But what happens next will be guided by the parameters of some theory or other. For, if we discover that the factories these people used to work in have closed down because the products they produced are no longer competitive, then we have the basis for invoking a theory.

Several options open up here. (1) On the assumption that markets once lost cannot be regained in an efficient way, find an alternative market that is underdeveloped and encourage its development. This can be accomplished in one of two ways: (a) encourage the market to function naturally, or (b) manipulate the market. If we choose (b), we can do something such as have the government provide low-interest loans for companies starting up in this area, provide tax breaks on capital investment, and so forth. (2) Our second general option with respect to the problem of getting people back to work operates on the assumption that products that have become noncompetitive can again be made competitive. Again, the same two alternatives are open: (a) on the assumption that the market works in certain ways, encourage the market, or (b) artificially manipulate the market. One way to manipulate the market is to increase tariffs. That has negative effects and we will not follow that option

any further. The determination that the effects of taking option 2b are negative, by the way, is not based on rational grounds. It is the function of following a certain ideological bent. But this is true of picking any option at this level of generality. All such choices of general approach are ideologically-based. Ideologies function in much the same way as theories do in guiding decisionmaking, but at a much more fundamental level. They are, furthermore, far more pervasive in their effects on our general outlook and on how we tend to interpret the events that impinge upon us.

We are now left with option 2a. Like 1a, 2a requires that we assume there is something called the market and that under optimal circumstances it will function in certain ways. If we choose either 1a or 2a, our recommendations will stress capital investment for either the development of new factories or the overhaul of old ones. That leads to the general solution: encourage the market so as to generate the required funds for capital investment. Thus, while the objective is to get people back to work, the problem to be solved is the development of capital investment. One solution to the problem is to encourage personal savings.

The point of this digression was to explore how theory determines the parameters for policy formation. Without the variables provided by the theories we could not even begin to think about the solution to the problem. We probably could not even recognize the problem as a problem of capital investment. To diagnose the problem as a problem of that kind requires a theory that allows us to conceptualize the cause of some specific set of events we see in front of us (such as lots of people being out of work) as ultimately, in the chain of reasoning, a function of lack of capital.

In the construction of a theory about the market and consumer behavior and the effect of government intervention, a number of key assumptions are made, the most important of which concern our beliefs about how the consumer will behave under given considerations. This implies a rather specific picture of the consumer. It also implies some standard of rational choice since even if the consumer cannot be expected to act rationally, on the account of rationality required by the theory, we, nevertheless, are supposed to formulate recommendations for changing behavior. If the consumer is expected to act irrationally from the point of view of our theory, we still must be in a position to predict the degree of deviation from our norm. This means that we have such a norm. But, more importantly, this also means that the consumer must act rationally even when deviating from our norm. That is, if the deviation is predictable and measurable, it means consumers act in accord with some select set of principles. The consumer has good reasons for choosing as he does.

Now since the theory is an idealization, it is essential, if the theory is to be at all useful, that we have knowledge of how the consumer acts. As just noted, this means we require two accounts. First, for purposes of making the idealization expressed by the theory make sense, we need some characterization of the ideal consumer as a decisionmaker. Then we need an additional account of the consumer based on empirical research to figure deviations from the norm as the norm is articulated in the theory. Thus we are led back from policy, through theory to some prior conceptualizations of economic agents as rational choosers. What we have also seen is that there may be two different accounts of "rational" at work here, the one according to the theory and the one actually employed by the agent. It is also the case that if the theory is to work at all, the consumer must be rational on some account or other. If not,

there will be no basis for making predictions and policy developers will be out of a job.

The Rational Consumer

Let us turn now to some presuppositions in the theory of rationality, particularly those pertaining to homo economicus. The following examination is by no means complete; it merely attempts to expose some of the presuppositions we carry with us when talk about consumer behavior and economic theory.

The most general account of the rationality of homo economicus can be summarized colloquially as "the rational man does not make book against himself." In other words, no rational agent intentionally chooses to lose. This opens the door to two different analyses. (1) The rational agent never intentionally picks an option that is less than optimal. This admits the possibility that in some situations there may be no optimal solution to a problem, thereby allowing the agent to succeed by minimizing losses. (2) The rational agent always seeks to win. On this view, the rational agent seeks to maximize his gains and minimize his losses. There are, of course, many variations on these two views, but generally homo economicus is conceived of as a game player with the basis for choices being the safeguard of self-interests.

There are many assumptions packed into the above account of rationality. We are going to look at some of these below. But before we do, I should like to speak to a prior question: is there a difference between choice and decision making? To make a decision is to choose one of several options. To follow the lead of existentialist Jean Paul Sartre, except in situations where physical necessity overrides all options, one is never in a position not to choose, for to choose not to do something is still to choose. Let us take the familiar case of deciding whether to harvest a crop today or tomorrow. Doing nothing about harvesting today is in effect a choice to harvest tomorrow. If the situation incorporates recognized consequences, then a choice is made.

As argued above, one is never in a position not to choose. Whether one exercises that option is another story. Sometimes one just proceeds through life without thinking much about what one is doing and whether one could do otherwise. But it does not follow that we could not think about, and in fact do, otherwise. The cost may be great, so great that we may decide not to do otherwise, but that does not mitigate against the possibility of thinking about it and choosing not to act.

Note also that on this view it does not follow that every choice is a rational choice, but only that we are never in a position not to choose. The caveat concerning physical necessity is intended to cover such exigencies as rain at harvest time, flat tractor tires, and so on.

If we turn now to decisionmaking, we can explore the idea that to decide is to select one of the available options, refusing to choose itself also being an option. If one is talking about rational decisionmaking, there are many theories of how one selects and on what basis. But, for ordinary, daily, run-of-the-mill decisionmaking, one makes a decision by choosing pair-wise A rather than B. In circumstances that warrant serious consideration, the situation becomes more complex. In those circumstances one takes pains to determine all available options; then the problem is to determine which

criteria to use to choose among them. Does one first narrow the set of options and then proceed pair-wise, or does one simply tally pros and cons for the entire list. Here, we also face problems about the distribution of weights, short- and long-term gains and losses, etc. At this point, problems of choosing will or rationally come to the fore.

An interesting empirical question concerns the degree to which consumer behavior reflects rational decisionmaking procedures for homo economicus versus ordinary daily types of decisionmaking. Except for rare purchases such as a house and the very mundane purchases for food, for the remaining discretionary budget, the consumer behaves irrationally on most accounts of rationality; there is no discernible pattern to decisionmaking. It has some serious consequences for linear decisionmaking models. In particular, if it is recognized that the more serious the objective, the more detailed the decisionmaking procedure, then it follows that not all decisions made without following some detailed model for homo economicus are necessarily irrational. There appears to be a sliding scale on which are correlated the importance of the decision with the detail in the decisionmaking procedure. If one does not use an elaborate decisionmaking model, it does not follow that the decision is irrational.

Presuppositions of Rational Decisionmaking

We are now in a position to consider some of the more important presuppositions of rational decisionmaking. Presupposition number one: choice is necessary. This presupposition is essential. Anything less would lead to undesirable results. We have already argued that not being able to choose is not an option. That leaves the question as to whether choice is necessary or merely possible. If it were a presupposition that choice merely be possible, then we could prescribe the parameters of decisionmaking in such a way as to preclude the possibility of ever being in a position to make a choice. There is a second consideration here. If choice were merely possible, then we would have to introduce a knowledge condition that could never be net, thereby making choice impossible. That is, if choosing were merely a possibility, it follows: (1) we would have to specify under which conditions choices are possible, and (2) we would have to know whether we were in those conditions before we could know if we could make a choice. There are two standard accounts of knowledge, on each of which you would never be in a position to make a choice. On the first account, knowledge is certainty. Surely you can not be certain of anything, hence you could never know if this was a circumstance in which you could make a decision. On the second account, knowledge is justified true belief. The issue turns here on the notion of justification: you would have to know you were justified, but that merely creates an infinite regress. In short, rational decisionmaking presupposes that choice is necessary.

Presupposition number two is that rational choice is possible. From the fact that choice is a necessary feature of homo economicus, it does not follow that every choice must be rational. Of course that depends on what we mean by "rational." But in order for applied economics to fall within the range of the sciences, it is essential that it be possible for economic decisions to be based on good reasons. This is not to say that they have been so based. Rather, it is only to argue that they must be capable of being so based or we pass from science to voodoo. Hence, it follows that an adequate account of 'rational' must be constructable. Many have been proposed. The ones we are

most concerned with are those articulated for homo economicus. They share a common feature: rational decisionmaking is a function of balancing needs, goals, and resources against each other and what is known. It is also assumed that all goals are generated by the individual, hence individuals are to be judged rational in terms of whether they select the options that maximize their goals.

This last feature is, to my point of view, the weakest part of the homo economicus model of rationality. It only makes sense to maximize one's goals if one has some determinate sense of how to rank-order them properly. That cannot be accomplished. No one has that kind of control over either their values or their goals. The major problem is that few of our goals are rationally or even consciously selected.

The only circumstances under which it would be appropriate to make maximizing goals a hallmark of rationality would be when all one's goals are self-generated. A self-generated goal would be, for example, the deliberate choice to devote one's life to solving a particular problem such as feeding the world. Some goals are consciously adopted goals that others have formulated, for example, to be a good person. These kinds of goals come equipped with lots of hidden assumptions and often prove troublesome for an individual when they conflict with the goals it has selected for itself.

The difficulty posed by goals formulated by others is that they do not automatically come equipped with a dictionary explaining what it means to achieve that goal. This is a partial function of the fact that goals are formulated in sets. When one party adopts one out of a set of goals formulated by a second party, it is likely that not only will the dictionary not be part of the adoption package, but the single goal may prove incoherent on its own. This is the sort of problem children face when they hold only a part of the value system their parents tried to inculcate and augment it with values and goals they have deliberately adopted. The resulting set of values most likely will be at best inconsistent, if not incoherent. It takes most of us a long time to sort all this out and some of us are bedeviled forever by the inconsistencies brought about by attempting to reconcile conflicting systems of values and goals. Finally, some goals are merely assumed as part of the baggage of being a member of a particular society. Such a goal might be to prove oneself a productive member of society. Integrating this third factor into an already complicated situation presents problems of yet another order of magnitude. Therefore, because all of an individual's goals are not self-generated in some pure sense, it would be unfair, if not irrational, to condemn a person for choosing not to maximize his or her goals.

Furthermore, because the goals one is expected to maximize are from diverse backgrounds and adhered to by an individual for different reasons, if there are reasons, our third presupposition (that the behavior of homo economicus must be predictable) is false. The argument for this conclusion is somewhat complicated, but consider the following.

First, there are different dimensions to rationality. We have already noted two of these: (1) the importance of the decision, and (2) the circumstances under which the decision is made. Depending on the degree of importance of the decision to be made, its long-term effects, etc., different standards of adequate reasoning will be brought to bear on the result by both the decision-maker and any outside observer. Likewise, for the dimension of circumstance. The decisionmaker's assessment of circumstances can explain and sometimes even justify a choice.

A third dimension of rationality is complexity, but a kind of complexity of the emergent kind. For us, that means that the complexity generates a new situation in an unpredictable manner. Biological life for example, has been described as an emergent property of straightforward chemical interactions. Similarly, increasingly complex societies emerge in unpredictable manners. What I am proposing is speculative, but in true philosophical tradition, I hope there is a germ of the right idea in it.

Preferences change over time; so do goals. And here we can introduce some reasons for distinguishing goals from preferences. Consider an 8-year-old who is given a choice between a candy bar and an evening of wild abandon with the person of his or her fantasies. Now consider that same person 20 years later and given the same choices. There is a good sense in which the goal remains the same: leisure. Now consider the case where the preference remains the same and the goals change. I still prefer pleasure to pain, but as an 8-year-old I associated pleasure with the achieving of one goal, being a firefighter, and as a 28-year old I associated it with being the world's best cook. These kinds of changes are commonplace.

The kind of change over time which presents major problems for general theories of rational homo economicus differs from the above and is complicated in two ways. First, goals, needs, resources, and knowledge are constantly assessed in the light of the success or failure of prior choices. While it may not always be conscious, we also are constantly reevaluating our understanding of our place in the world and our knowledge of the world. This reevaluation is the heart of the common-sense approach to living. It may even be the heart of rationality. For the refusal to reassess a position in light of new information is the best example of irrationality. But there is more. If goals are not mere preferences, and if they are constantly under assessment regardless of how acquired, then rationality theory must accommodate a continual sorting out of goals over time, independent of considerations of circumstance and relative importance. Assessment of goals normally takes place in a pair-wise, just as rational decisionmaking occurs. Thus, goals such as career, quality of life, role in society, self image, role as husband and father, and so on, are assessed by comparing either some other state within the same goal-context, or by comparing two goals.

Finally, the sum total of one's goals implies something else. Rationality does not allow the rank-ordering of goals with a checkoff as they are achieved. In fact, that is not part of our view of rationality at all. The complexity of the interaction among goals is a continuous function.

The rational person is constantly juggling goals and reestablishing priorities in the light of changing circumstances. And when the juggling becomes too frantic, at a time not predictable, a different control takes over. Part of the common-sense assessment procedure is simplifying when things get too complicated. How is it to be accomplished? To introduce consistency into a set of options, one derives a subset of goals which form the basis for future decisionmaking. If we abide by the results of these abstract deliberations, the consequence is a change in life-style. This is an emergent feature of the complexity factor.

Is there a way to accommodate these kinds of changes with respect to the complexity of goals and their consequences in a theory of rational decision-making? Probably not. Such changes are instituted when a threshold of complexity is met. That threshold will differ for each person. Nevertheless,

unless some account is taken of the fact that life-style changes do occur, then no adequate planning can be forthcoming. For such changes have effects on the market. Furthermore, as the complexity of contemporary society increases and as the rate of change in society continues to advance, the rate of systematic reevaluation leading to life-style changes will also increase dramatically. So as our final presupposition of rational decisionmaking we have the false belief that all changes relevant to rational choice are market based and are systematically assessable. They are not.

Conclusion

This paper seeks: (1) to provide an overview of the kinds of things philosophers do and how they do them, and (2) to raise some questions about choice and rationality, which are interesting from a philosophical point of view, but which also have a bearing on economic theory.

Solutions are not proposed, even when available. Furthermore, before solutions are set out, ramifications with respect to the history of thought are needed. In other words, it is necessary to know how it all hangs together.

References

1. Churchman, C. W., Prediction and Optimal Decision: Philosophical Issues of a Science of Values. Englewood Cliffs: Prentice-Hall, Inc., 1961.

2. Glymore, C., Theory and Evidence. Princeton: Princeton University Press, 1980.

3. Hausman, D. M., Capital, Profit and Prices: An Essay in the Philosophy of Economics. New York: Columbia University Press, 1981.

4. Hempel, C., Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. New York: The Fress Press, 1981.

5. Pitt, J., Philosophy in Economics. Dordrecht: Reidel, 1981.

6. Plato, Republic. Translated by G.E.A. Grube. Indianapolis: Hackett Publishing Co., 1974.

7. Rosenberg, A., Microeconomic Laws: A Philosophical Analysis. Pittsburgh: University of Pittsburgh Press, 1976.

8. Sassower, R., Philosophy of Economics: A Critique of Demarcation. New York: University Press of America, 1985.

9. Sellars, W., Science, Perception, and Reality. London: Routledge and Kegan Paul, 1968.

COMPARATIVE EFFICIENCY JUDGMENTS, ECONOMICS, AND POLITICAL PHILOSOPHY

by Allen Buchanan

The aim of this essay is to make clear the extremely important role that intersystemic efficiency comparisons play in political philosophy and to articulate the hitherto unappreciated difficulties involved in making such comparisons. This topic is only one element of a more ambitious effort (1) to bridge the gap between two different approaches to evaluating the market. 1/ Economists have tended to evaluate the market solely in terms of efficiency, neglecting ethical issues. Ethicists (including philosophers and normative political theorists) have characteristically disdained considerations of efficiency (when they have thought about them at all), while concentrating on the moral assessment of the market, most recently in terms of its failure or success in satisfying the requirements of justice or equity.

This sharp division of labor between ethicists and economists is untenable and rests upon the failure of both groups to reflect critically upon the presuppositions that underlie their respective approaches to the market. The adoption of the allegedly purely technical notion of efficiency that prevails in the economics literature itself rests upon controversial moral assumptions. Conversely, some of the most influential ethical arguments for and against the market rest on unarticulated assumptions about the efficiency or inefficiency of the market or rival systems. A major theme of this paper is the complex and often inexplicit interplay between efficiency assessments and moral judgments. Another is the precariousness of attempts to compare whole systems on grounds of efficiency. The charge of relativism has long been leveled at attempts to make comparative ethical judgments about different social systems; the problem of relativism for judgments of comparative efficiency turns out to be less daunting.

Intersystemic Efficiency Comparisons and the Paretian Principles

An intersystemic efficiency comparison means here a judgment that ranks two (or more) economic systems as to efficiency, using the Pareto Optimality Principle or the Pareto Superiority Principle in their most inclusive forms. The Pareto Optimality Principle and the Pareto Superiority Principle in their most inclusive form may be stated as follows.

> Pareto Optimality: A state, S^1, of a given system, S, is Pareto Optimal if and only if there is no feasible alternative state of that system, S^2, in which at least one person is

---

Allen Buchanan is a Professor of philosophy, University of Arizona, Tucson, AZ.

1/ Underscored numbers in parentheses refer to sources listed in the References at the end of this article.

better off and no one is worse off.

Pareto Superiority: A state, S^1, of a given system is Pareto Superior to another state of that system, S^2, if and only if there is at least one person who is better off in S^1 than in S^2 and no one is worse off in S^1 than in S^2.

The Pareto Optimality and Pareto Superiority Principles may be formulated more narrowly to refer exclusively to distributional states. In other words, taking certain stock of consumer goods (including services) as given, we can ask whether a particular distribution of them, D, among the persons in the system is such that there is no alternative way of distributing those goods, D^1, among those individuals, that would make at least one individual better off without making anyone worse off (Distributive Pareto Optimality). And we can ask whether a change from one distribution (D^1) of a given stock of consumer goods among a set of persons to another distribution (D^2) would improve the condition of someone without worsening the condition of anyone else (Distributive Pareto Superiority).

On the other hand, the Principles of Pareto Optimality and Pareto Superiority can be applied not to distributions of consumer goods, but instead, exclusively to allocation of resources for producing goods. An allocation of productive resources is Pareto Optimal if and only if there is no (technically possible) alternative allocation that would produce more of at least one good without producing less of some other good (Productive Pareto Optimality). Similarly, if an allocational state, A, would produce more of a least one good than an allocational state, A^1, but without producing less of any other good than A^1, then allocational state A is Productively Pareto Superior to allocational state A^1. The more inclusive formulation of the Pareto Optimality and Pareto Superiority Principles stated earlier are broad enough to combine allocational and distributional assessments if $'S^2$ and $'S^1$ are understood as referring to allocational distributional states. An allocational-distributional state consists of both a given allocation of productive resources and a coexisting distribution of consumer goods (and services).

Economists sometimes distinguish yet another type of efficiency. A system is said to be Aggregatively Efficient if it employs all available productive resources; Aggregatively Inefficient if it does not. Critics of private property market economic systems often argue that that system is Aggregatively Inefficient because they regularly suffer from unemployment, which is a failure to utilize available labor. However, Aggregative Efficiency as it stands is an incomplete evaluative criterion for two reasons. First, what is important for the evaluation of a system is not whether a productive resource is being used but whether it is being used in an efficient way. The fact that certain resources are not being used need not constitute a significant defect of a system if they are not needed or if their use would be unproductive or inefficient. For these reasons it seems best to subsume Aggregative Efficiency under Productive Pareto Optimality. Thus, unemployment (of potentially productive workers) would result in Productive Pareto Sub-Optimality only if putting the unemployed to work would result in the production of greater quantities of some goods without reducing the quantities produced of other goods. Further, even if certain idle resources could be used to produce a particular item, this will be of little consequence if the item that would be produced would not in fact increase anyone's well-being, either because the thing is not desired by anyone to whom it is distributed, or because it is not distributed to anyone who desires it, or because the resource is worth more than the product it would yield. The main limitation on the usefulness of Aggregative Efficiency as a criterion for assessing social systems is that it

does not include the appropriate connections between resources, productivity, and benefits from what is produced.

Notice that the notion of Productive Pareto Optimality suffers from a similar, though more limited shortcoming. By focusing only on the quantity of particular kinds of outputs attainable, Productive Pareto Optimality fails to engage the root idea behind the Pareto Optimality and Pareto Superiority Principles: the effect of the particular form of social organization in question on the well-being of the individuals involved. For clearly a system would be Productively Pareto Superior to another without making anyone better off if the additional goods it produced made no contribution to anyone's well-being. For instance, the fact that system S can produce more nuclear weapons than system S^1 without reducing its production of other goods does little to establish the superior efficiency of system S if more nuclear weapons are of no benefit, unless we are willing to separate the notion of efficiency from that of well-being entirely. This point will become important later when we examine attempts to compare the efficiency of rival systems by comparing their productivity.

A related but distinct conception of productive efficiency is what is commonly called "productivity," that is, the ratio of outputs to inputs. One is more productive than another in this sense if it produces a greater output than another, granted the same input. This second conception of productive efficiency differs from Productive Pareto Optimality insofar as it measures productivity relative to a constant resource base. The Productive Pareto Optimality Principle in contrast does not assume equal inputs.

Productivity, that is, productive efficiency as maximization of outputs relative to inputs, like Productive Pareto Optimality, has no necessary connection with well-being. A system S may produce more of a good G than a system S^1, using the same quantity of inputs I, yet no one in S^1 will be better off in S than in S^1 unless the extra quantity of G brings an additional benefit to someone.

To compare two systems on grounds of productivity, one must select some particular type of output or set of types of outputs (for example, steel and bread) that are common to both systems and compare the amounts of these produced in the two systems to some constant quantity of a type of input (or set of types of inputs) which is also common to the two systems. More important, it is also necessary to assume that the particular products (outputs) and resources (inputs) that provide the basis for the comparison are equally valuable to individuals in the two systems, if productivity comparisons are to be reliable indicators of comparative efficiency, and if we are to assume that efficiency has some relationship to how well off people are. For example, the fact that one system produces more guns per unit of iron than another is of limited interest if individuals in the two systems place sharply different values on guns. To assume that the products or resources identified in productivity comparisons have the same or roughly the same values for individuals in two different systems is to assume interpersonal utility comparisons; comparisons of how well off one individual is relative to how well off another is. This, as we shall see, is a problem, since one of the main motivations for using the Paretian concept of efficiency was to avoid having to make interpersonal utility comparisons.

For the same reason that productivity is inadequate for overall efficiency assessments, neither the growth rate, nor the rate of capital accumulation is by itself a satisfactory measure of a system's efficiency, unless we divorce the concept of efficiency from that of well-being. A system with a higher rate of accumulation or a higher growth rate might not make people better off if it suffered from an inefficient distribution system, for example.

Among the efficiency principles discussed, the Pareto Optimality and Pareto Superiority Principles in their broadest forms appear to provide the most comprehensive tools for assessing a system's efficiency, since the notion of a social state they employ is inclusive enough to take into account the way productive resources are allocated, the way production is organized, and the distribution of consumer goods so far as all of these affect how well off individuals are. However, since the Pareto Optimality and Pareto Superiority Principles focus only on the impact of these factors upon individual well-being, not upon the relative contribution that each factor makes to individual well-being, a judgment that a particular system is in a Pareto Sub-Optimal state or that one social state is Pareto Inferior to another by itself tells us nothing about whether the source of inefficiency lies in the allocation of productive resources among producers, the ways in which production is actually organized, or in the mechanism for distributing what is produced. In what follows, the relevant notions of efficiency will most often be that of Pareto Optimality and Pareto Superiority in their inclusive forms, as covering both distribution and allocation, though in some instances the more specific Distributional or Productive Pareto Optimality and Superiority Principles may be invoked.

The wide acceptance of the Pareto Optimality and Pareto Superiority Principles is usually attributed at least in part to the fact that they provide a way of assessing social states that does not require interpersonal utility comparisons. While there is considerable, though not unanimous, agreement that interpersonal utility comparisons cannot be made, there has been sharp difference of opinion as to why they cannot be made. Some have contended that such comparisons are impossible because one person cannot know enough about another person's "inner state" of satisfaction or pleasure to be able to compare it quantitatively with his own (9).

There are two serious problems with this reason for concluding that interpersonal utility comparisons cannot be made. First, it assumes an excessively narrow conception of satisfaction (or pleasure or utility), one that identifies it as a psychological state and, more particularly, an affective state of awareness or feeling (3, 7). Although it is no doubt true that the terms "pleasure" and "satisfaction" are sometimes used to refer to such a state of awareness, we also often speak of activities (such as playing tennis or producing a painting) as pleasurable or satisfying, even if they are not accompanied by a particular feeling or psychological state of pleasure or satisfaction that is distinct from the activity itself.

More important, this first argument for the impossibility of interpersonal utility comparisons rests upon a now-discredited philosophy of mind which assumes that one can only know of the existence or character of a mental state by introspection (or "direct" acquaintance). Since one can only introspect one's own mental states, no one can know of the existence or character of anyone else's mental states of any kind, not just states of pleasure or satisfaction. However, once a broader view of knowledge is

adopted, and the assumption that mental states are known only by introspection is abandoned, this first argument for the impossibility of interpersonal utility comparisons loses its force. In particular, a belief in the existence of mental states is justified by the fact that postulating their existence is required for developing certain behavior. On this account, talk about mental states, and claims to knowledge of them, are warranted by their explanatory power, within the context of a theory.

The second argument against interpersonal utility comparisons is the lack of a common zero point or baseline from which to measure different individuals' utilities, and lack of a common unit of measurement. Even if a utility scale can be constructed for each individual by recording the choices he or she makes among various goods or options, observation of different individuals' revealed preferences provides no basis for relating their utility scales to one another. This second difficulty, unlike the first, depends neither upon an unduly narrow conception of utility as a mental state nor upon the dubious epistemological view that mental states are radically private.

Against this background, the two Paretian Principles (the Pareto Optimality and Pareto Superiority Principles) are usually seen as second-best alternatives to utilitarianism, the assumption being that assessing social states according to the overall utility they produce would be preferable, were it not for the unfortunate fact that interpersonal utility comparisons cannot be made. The Paretian principles avoid interpersonal utility comparisons by requiring only that we be able to determine whether each individual is better off or worse off relative to his own former condition.

The Role of Intersystemic Efficiency Comparisons in Political Philosophy

Political philosophers have often observed that a social system may be efficient (in the Paretian sense) and yet unjust. For example, a system in which most have nothing and a few have everything may be Pareto Optimal, since improving the condition of the unfortunate majority may require worsening the condition of the privileged minority. Nevertheless, even though they have acknowledged that efficiency is not a sufficient condition for a just or good society, political philosophers acress the entire ideological spectrum have found it necessary to assert that the systems they favor are efficient.

For utilitarian philosophers, the need to rank alternative systems on grounds of efficiency is obvious enough: a system that is efficient produces more utility than one which is not. Thus even if the problem of interpersonal utility comparisons prevents the utilitarian political philosopher from determining which system among the feasible alternatives maximizes utility, he can still determine which of two alternatives produces more utility than the other if he knows that one is more efficient than the other.

Intersystemic efficiency comparisons are important for nonutilitarian political philosophy as well. In chapter V of The Second Treatise, Locke attempts to justify the current unequal distribution of wealth by arguing that the private property system which produces it makes everyone better off than he or she would be in a condition of equality. Locke (and after him Adam Smith) boldly claims, without clarification and without argument, that even the poorest person in the market system of his day is better off than the richest member of the nonmarket system in which American Indians were thought to live (5).

Marx, as well as later Marxists including G.A. Cohen, also stake a great deal on the ability to make intersystemic efficiency comparisons. They contend that communism will be more productively efficient than capitalism and that it will also distribute what is produced more efficiently (2). For Marx and at least for more or less orthodox contemporary Marxists, this prediction that communism will be more efficient plays a crucial role in the account of what life in communism will be like and of how it will be superior to life in other types of societies. In particular, Marx seems to believe that the abundance that this greater efficiency will make possible is an indispensable precondition for the fullest development of every individual within a context of harmonious social relations (6). Yet neither Marx nor later Marxists have taken seriously the problem of understanding, much less justifying, comparisons between quite different systems as to efficiency.

For Locke the alleged greater efficiency of a system of private property rights is cited to help defend that system against the charge that it produces unjust inequalities. For Marx, the alleged greater efficiency of communism serves both as a key premise in the critique of capitalism and as reason for preferring communism. In Rawls, intersystemic efficiency comparisons play a less foundational, but still vital role (8).

Unlike Locke and Marx, Rawls does not advance an intersystemic efficiency comparison and then use it as a premise in criticizing one system and supporting another (8). Instead, Rawls's theory is itself scrupulously agnostic about whether a system with private property or a socialist system would be more efficient and hence, would provide greater resources for improving the prospects of the worst off. Yet, once we descend from the realm of pure theory and attempt to implement Rawls's Difference Principle, intersystemic efficiency comparisons become necessary. We can determine what basic institutional arrangements Rawlsian justice requires only if we can compare alternative systems on grounds of efficiency. Unless such comparisons can be made, Rawls's theory cannot even tell us whether the just society will be capitalistic or socialistic. It should be clear, then, that intersystemic efficiency comparisons have played and continue to play a vital role in political philosophy.

The Barriers to Intersystemic Efficiency Comparisons

But there are rather severe obstacles to making intersystemic efficiency comparisons. The depth of the problem can best be appreciated by examining several alternative proposals for grounding such comparisons and uncovering their problematic presuppositions. We noted earlier that there is broad agreement that efficiency is to be understood as Pareto Optimality and that the concept of Pareto Superiority is to be used for making comparative efficiency judgments. However, the Pareto Superiority Principle is designed to permit comparisons between different states of the same system, as to how those different states affect the well-being of the same set of individuals. It is not designed to compare different systems, which contain different individuals.

Nonetheless, one might attempt to extend the Pareto Superiority Principle to intersystemic efficiency comparisons by developing the concept of a representative (or typical) state of a given system. The idea would be to pick out a representative state of another system and then to compare the systems by determining whether the representative state of one was Pareto Superior to the representative state of the other. But developing the notion of a representative state adequate for such comparisons may be difficult.

Even if the concept of a representative state of a system can be adequately developed, however, there is a much more serious problem. Suppose that we are attempting to compare the Soviet system with the American system by applying the Pareto Superiority Principle to their representative states. Presumably we are to ask whether at least one individual in the representative state of one system is better off than anyone in the representative state of the other system while none on the representative state of the first system are worse off than anyone in the representative of the second system. Since the different systems contain different groups of people, answering this question requires us to make interpersonal utility comparisons to compare one individual's well-being with that of another. But one of the main reasons for adopting the Pareto Superiority Principle in the first place, it must be recalled, was that its use avoids interpersonal utility comparisons. To that extent, this first approach to intersystemic efficiency comparisons seems unacceptable.

A second strategy attempts to reduce the problem of comparing two different types of systems to that of comparing two different states of the same system that has undergone a change from being a system of one type to being a system of another type. Suppose we wish to use the Pareto Superiority Principle to compare systems of type A (for example, market systems) with systems of type B (for example, nonmarket socialists systems). We focus on a particular instance of a type A system, call it system 1. We then perform a thought experiment in which we imagine systems 1 undergoing a transformation from being a type A system (in representative state R^b), for example, a transition from capitalism to nonmarket socialism. We then compare R^a (the representative state of 1 when 1 has become a type B system). This way of making intersystemic efficiency comparisons attempts to avoid interpersonal utility comparisons by focusing on only one set of individuals (and asking whether at least one of those individuals would be better off in the one type of system than he would be in the other type, while no one would be worse off in the first system).

The second approach, however, is also vulnerable to objection. It rests upon two dubious assumptions. The first is that the transformation from one type of social system to another does not radically alter the interests of the individuals in question. The second assumption is that an individual can be shown to be better off in one state than in another in which interests in the two states differ radically.

The first assumption would be unproblematic if an individual's interests (that is, what contributes to his well-being), were given and fixed independent of the type of social organization in which he lives. 'Interests' may be construed either subjectively or objectively. According to the purely subjective notion of interests, something is in a person's interest only if he or she is interested in it. The purely subjective notion of interests identifies interests with preferences. Clearly, if interests are understood in this purely subjective way, then the sort of society in which one lives plays a very important role in determining what is in one's interests, because at least some of one's preferences develop as a member of a particular society.

At the other extreme, 'interests' may be construed in a purely objective way. What is in one's interest (that which if attained or realized will contribute to one's well-being) may be wholly divorced from what one prefers (that in which one takes interest). At least when applied to normal individuals with adequately developed powers of choice and deliberation, this purely objective view of interests has little to commend it. For although an individual may

sometimes, because of ignorance or weakness of the will, fail to prefer or take an interest in something that is conducive to his or her well-being, it is plausible to maintain that an individual's well-being is generally based on his or her actual preferences.

Nonetheless, an individual's interests may still be partly objective in the following sense. There may be some things, such as an adequate minimum of food, shelter, and rest, in which virtually all people do in fact take an interest and there may be some states of affairs, such as the existence of some form of social cooperation, which are almost universally necessary for attaining these things. The interests in these goods and in the conditions for attaining them may properly be called objective interests. However, in any circumstances beyond those of mere subsistence, most individuals will regard these objective goods (such as food, shelter, and rest) not as ultimate ends, but as instrumentally valuable for the attainment of other things they prefer. And which ultimate ends an individual pursues, as well as how he or she relates them one to another in order of priority, may be greatly influenced by the social environment in which his scheme of goals is formulated. Hence, even if what is conducive to an individual's well-being is in part determined by objective interests, it is a mistake to assume that the transformation from one type of social system to another does not radically alter the interests of the individuals in question.

The second assumption, that we can reliably determine whether or not an individual is better off in one state than another even though his interests in the two states are radically different, may be challenged in either of two ways. The first, more extreme of these is to argue that if the change in an individual's interests wrought by a transformation of the social system is profound enough, then the conditions for the individual's personal identity are undercut. The assumption here is that it is the character of one's most fundamental interests that makes one the particular person one is. If the individual whom we called Jones in the type A system is a different person from the individual we call Jones in the type B system, then our thought experiment is of no avail since it is comparing the well-being of two different Joneses. But recall, the purpose of the thought experiment was to develop a way of making intersystemic efficiency comparisons that did not involve interpersonal utility comparisons.

This line of argument, however, is open to a serious objection. According to a plausible theory of personal identity developed by John Perry, a person can survive even rather drastic changes in his interests (7). This theory distinguishes between values, preferences, interests, and other features of what we commonly call personality, on the one hand, and the underlying properties of processes whose existence is postulated to account for the rather dramatic changes in the constituents of personality that may occur in the normal course of maturation and development over the individual's lifetime. The theory holds that so long as there is continuity in the substratum properties or processes, the same person persists, even if there are radical discontinuities in the surface psychological features. So if this view is correct, the same person might continue to exist even if a transformation of the social system produced a profound change in his or her interests.

The second way of challenging the assumption that we can reliably determine whether an individual is better off in one society than another even though his interests are radically different in the two societies does not depend on the claim that the conditions for personal identity have been undercut. The idea,

rather, is that we cannot reliably judge that Jones is better off in society 1 than in society 2 unless there is a substantial common core of interests that Jones has in both societies.

It might be thought that the objective interests discussed earlier, the interest in food, shelter, and rest, suffice to serve as this common core. It is true that we may be confident in judging that Jones would be better off in society 1 if these basic interests were satisfied in society 1 but not satisfied in society 2. However, once we are faced with the task of comparing Jones's well-being in two societies that both satisfy these basic interests, our confidence wanes if the differences in the other interests Jones has in the two states are profound. Yet for the comparisons in which we are most interested (for example, comparisons between market systems and socialist systems) the more plausible versions of both will presumably provide for basic interests.

We could confidently make the comparative judgments in question if in doing so we were simply objectively measuring different quantities of the same "stuff" in the two societies; magnitudes of pleasure, for example, where pleasure is thought of as a psychological state or a property of a psychological state. We have already seen, however, that even if well-being could be reduced to pleasure, there is no reason to restrict pleasure to a psychological state that could be quantitatively evaluated.

To avoid the problematic metaphysical assumption that in comparing Jones' well-being in two societies in which his interests are radically different we are objectively measuring different quantities of the same "stuff," we might instead attempt to rely upon Jones' own judgment as to his comparative well-being. One can certainly imagine an individual who lived through, and internalized, such a dramatic transformation of the social system himself rejecting any attempt to compare his state of well-being before and after the transformation. But even if he confidently claimed to be better off now than he was then, it is not obvious that we should consider his judgment well grounded. For if he has truly internalized the values of the new social system, he will no longer be capable of accurately appreciating the value of the life he used to live. If the discontinuity between his interests in society 1 and his interests in society 2 were profound, we might conclude that reliable comparisons of his states of well-being in the two societies were impossible, regardless of his confidence in his ability to make them.

We are most confident about comparisons of well-being of an individual in different states if there is a substantial core of interests common to both states. In fact, our confidence is greatest where we assume that the interests relevant to the comparison are the same, since we typically limit our comparisons to some subset of the individual's interests. In other words, we more frequently make local and limited intrapersonal well-being comparisons rather than global, all-inclusive ones.

We are usually more interested in knowing whether some change in an individual's situation has furthered or hindered certain of his or her interests rather than whether it has increased or decreased his well-being as such. These central cases of comparisons differ from the problem of determining whether, for instance, the individual we now call Lee is happier living as a businessman in Hong Kong than he was living in the Stone Age culture into which he was born and in which he lived for 20 years on a remote island in the Philippines. Even if our criteria for personal identity allow us to say that the same person persisted through the change from one type of

society to another, the change in the individual's interests may be so great that the problem is functionally equivalent to that of making interpersonal utility comparisons, strictly speaking.

Another, no less serious difficulty is that even if the Pareto Superiority Principle could be extended to intersystemic comparisons without resurrecting the problem of interpersonal utility comparisons it was invoked to avoid, the principle cannot provide a ranking of the types of systems we are most concerned to compare. Recall how stringent the Pareto Superiority Principle is: a state is Pareto Superior to another only if no one in the second state is worse off than anyone in the first. Surely when a socialist claims that socialism is more efficient than capitalism, he is not necessarily denying that some individual (for example, the richest capitalists) will be disadvantaged by a transition to socialism. Nor do those who contend that capitalism is a more efficient system than socialism thereby commit themselves to defending the equally implausible thesis that transforming a socialist system into a capitalist system would worsen no one's condition, not even that of the most privileged party bureaucrat. Just as the Pareto Optimality Principle tends to be useless when applied to nontrivial policy decisions within a given society, because in the difficult cases each option will affect someone adversely, so the Pareto Superiority Principle seems just as unhelpful if we attempt to use it to evaluate a transformation of the entire social order. To imagine that it would be otherwise is to be naively optimistic about the problem of harmonizing interests.

A third proposal for making intersystemic efficiency comparisons is designed to avoid the problem of changes in the individual's interests by abstracting entirely from the content of the individual's preferences. The suggestion is that we can compare different systems by comparing the preference satisfaction ratios of the same individuals in the two systems. The preference-satisfaction ratio, S/P for an individual, i, in a system, A, is the proportion of i's satisfied preferences to his total set of preferences. Thus we may compare two systems 1 and 2 as follows: if all (or most) individuals would have a higher S/P system in 1 than in 2, then 1 is the more efficient system. This way of making intersystemic efficiency comparisons, like the second formulation of the Pareto Optimality/Pareto Superior Principles stated above (in terms of preferences), assumes that the satisfaction of an individual's preferences either constitutes or at least accurately corresponds to his well-being. The virtue of this third approach is that we can compare S/P's even if there are radical changes in the individual's interests.

There is, however, a very serious limitation on the usefulness of the notion of preference-satisfaction ratios in making intersystemic efficiency comparisons. Suppose, for example, that in society 1, individual i has 100 preferences, and that in society 2, 10 of those preferences are satisfied. Suppose that in society 2, individual i would have only five preferences, but four of the five preferences would be satisfied in society 2. The S/P for individual i in society 1 is 1:10 (0.1), while in society 2, i's S/P is 4:5 (0.8). But surely it does not follow that i is better off in society 2 than in society 1.

For one thing, the quality of satisfaction of the preferences that are satisfied in 1 might be higher than the quality of satisfaction of the preferences satisfied in 2. It might be possible, for example, to reduce an individual's set of preferences to a very basic, easily satisfiable minimum, through psychosurgery, mind-altering drugs, or brainwashing techniques. But

even if a higher proportion of these limited preferences were satisfied, it would not follow that the individual would be better off. Comparisons of preference-satisfaction ratios in different systems can perhaps provide a rough indication of relative efficiency among systems, but the tenuousness of the connection between preference-satisfaction ratios and individual well-being restricts its usefulness considerably.

Yet a fourth approach would be to abandon the Pareto Superiority Principle altogether, acknowledging that when used in intersystemic comparisons it provides no ranking of interestingly different systems. Instead, only the Pareto Optimality Principle would be employed. One might argue that one system is more efficient than another if the first more frequently attains Pareto Optimal outcomes than the second (or if its representative state is more frequently Pareto optimal). The virtue of this approach is that it does avoid interpersonal utility comparisons.

The simple but fundamental problem with this approach is that it is unlikely that any real-world systems, or at least any of the real-world systems we are interested in comparing, ever reaches a Pareto Optimal state. In every case, some if not all of the ideal conditions that guarantee Pareto Optimality are imperfectly satisfied and this is true regardless of whether the system in question is a capitalistic (private property) market system, a centralized socialist system (such as the USSR), or a mixed system with some centralized planning plus decentralized decisionmaking among competing firms (as in Yugoslavia). The principle of Pareto Optimality is a binary evaluative criterion: a state is either Pareto Optimal or it is not. There is no such thing as being more or less Pareto Optimal. Efficiency in this sense does not admit of degrees nor, strictly speaking, of any comparison at all.

In general there is no uncontroversial, reliable, and even reasonably precise method for determining how much overall benefit is lost in a given system, at a particular time, due to failure to achieve Pareto Optimality, nor for determining the relative contributions to lost overall benefit that various sources of inefficiency make. Moreover, any attempt to determine how much loss of possible benefit can be attributed to a particular source of inefficiency, or to determine how much overall loss of possible benefit the system suffers as a result of various inefficiencies would require interpersonal utility comparisons, since both of these determinations would involve summing up benefits (utility) across different individuals.

It is important not to underestimate the significance of these points for arguments for or against a particular type of socioeconomic system on grounds of efficiency. It is not simply that there are practical difficulties (lack of information, for example) that hinder intersystemic efficiency comparisons. Rather, there is a sense in which there is not even a theory of intersystemic efficiency comparisons in the same sense in which there is a theory of intrasystemic efficiency comparisons. All arguments concerning the relative efficiency of rival systems of social organization would appear to have an irreducibly intuitive element, though this is not to say that statements about the relative efficiency of various systems are arbitrary. At best we can identify various sources of inefficiency that occur in the rival system, and then try to make some rather rough and ready estimate of how serious these inefficiencies are.

No attempt was made here to prove the impossibility of coherent, well-supported comparisons of different types of economic systems on grounds of efficiency. However, efficiency arguments for or against the market depend ultimately upon

such comparisons and, there are very serious, and often unappreciated, obstacles to making them, especially in cases in which the systems to be compared differ dramatically.

There is an interesting argument which may to some extent mitigate this skeptical conclusion. We have a theory of the ideal market system that is powerful enough to allow us to prove that an ideal market system achieves Pareto Optimality. However, we do not have a theory of an ideal nonmarket system on the basis of which to prove that this ideal nonmarket system achieves Pareto Optimality. Therefore, even though no actual market system satisfies the conditions of the ideal market system, the reasonable presumption is that actual market systems are more efficient than actual nonmarket systems. Let us call this the "some theory is better than no theory argument."

David Friedman advances this argument and offers an analogy to illuminate it (4). Suppose you wish to fire a cannon at a distant target. You have had an elementary physics course that includes a theory of the trajectory of a ideal projectile: the path that a point mass will traverse in a vacuum in a uniform gravitational field. Of course the cannon ball is not a point mass (it has extension), and the path it follows will not be in a vacuum (it travels through the air, with friction), and the earth's gravitational field is not uniform. But nonetheless it surely would be more reasonable to aim the cannon at the angle arrived at by calculating according to the ideal theory than to choose an angle at random! Similarly, our best shot at efficiency is to choose a system that at least approximates the ideal of efficiency; a system for which there is a theory capable of generating an efficiency theorem, rather than one for which there is no such theory. Thus the burden of argument must lie on the advocate of a nonmarket system to show that such a system is more efficient. And in the absence of a developed theory of the nonmarket system, that burden of argument cannot be successfully borne. So we must conclude that a market system is more efficient.

The argument can be strengthened. One advantage of choosing a system for which there is a theory capable of generating an efficiency theorem is that the theory will provide us with guidance as to how we can modify the nonideal system so as to bring it closer to the ideal conditions that guarantee efficiency. The theory of the ideal market system explains which features of nonideal market systems (monopoly, and so on.) produce inefficiencies and hence tells us how to bring our nonideal market system closer to the ideal of efficiency. If no comparably powerful theory is available for the nonmarket system, then we know that it is highly probable that the system is inefficient, but we lack theoretical guidance for improving its efficiency. Again, some theory seems better than no theory, even if reality is only a distant approximation of the system described in the theory.

This argument might conclusively show that a nonideal market system is likely to be more efficient than a nonmarket alternative if (a) there was in fact no theory of nonmarket social organization and if (b) there was no empirical evidence to show that nonmarket systems better approximate the ideal of efficiency than a random allocation would. However, neither of these assumptions seems to be justified. There is a theory of the planned socialist, nonmarket economy powerful enough to generate an efficiency theorem. It has been demonstrated mathematically by Pareto and Barone that such a system can duplicate the Pareto Optimal outcomes of the ideal market system. Further, the recent economic records of largely nonmarket systems such as the Soviet Union, though far from admirable, do not reveal utter economic chaos or even lack of growth, and from this it seems reasonable to conclude that these systems at

least approximate efficiency more closely than a purely random allocation would.

So even if it would be more reasonable to choose a nonideal market system over a random "nonsystem," we cannot confidently conclude that any actual market system, no matter how imperfect, is preferable, on grounds of efficiency, to all nonmarket systems. Further, even if we could draw this conclusion, the argument would still not show that even greater efficiency cannot be achieved by selective intervention to minimize externalities and other market imperfections, or to redistribute income the market process yields, by intervening in nonideal market systems. In other words, we would still be left with the problem of determining whether a more regulated or a less regulated, nonideal market would be more efficient. The conclusion to be drawn is not that the notion of efficiency is useless, of course. But it does seem clear that we are on much firmer ground in asking whether a relatively minor, incremental change (which would not produce a new system but only a modification of an old one) would be more efficient, than we are when we facilely claim that one system is more efficient than a quite different system. However, this counsel of caution will seem to some to have unfortunately conservative implications: efficiency arguments for radical change, as opposed to reform, appear to be shakey if not impossible.

Many who advocate the use of the Paretian concept of efficiency for comparing systems do so because they believe attempts to compare systems on ethical grounds are afflicted by an irreducible relativism: it seems to them that all such ethical judgments are valid only within the assumptions of a particular system of ethical principles and that there are no rational grounds for choosing one such system over another. They have failed to notice that the problem of relativism is no less serious for comparative efficiency judgments than for comparative ethical judgments.

References

1. Buchanan, A. E., Ethics, Efficiency and the Market, Totowa, N.J.: Rowman & Allenheld Publishers, 1985.

2. Cohen, G. A., "Labor, Leisure, and a Distinctive Contradiction of Advanced Capitalism," Markets and Morals. G. Dworkin, G. Bermant, and P.G. Brown, Eds. Washington, DC: Hemisphere Publishing Co., 1977, pp. 107-36

3. Cowan, J. L. The Concept of Pleasure. London: Macmillan, 1968.

4. Friedman, D., unpublished paper presented at a conference on "Health Care and the Market," Rice University, November 1985, sponsored by the Liberty Fund, Inc.

5. Locke, J., Second Treatise of Civil Government. 2d rev. ed., Indianapolis: Bobbs-Merrill, 1952.

6. McLellan, D., ed., Karl Marx: Selected Writings. Oxford: Oxford University Press, 1977.

7. Perry, J., ed., Personal Identity. Berkeley: University of California Press, 1975.

8. Rawls, J., A Theory of Justice. Cambridge, Mass: Harvard University Press, 1971.

9. Robbins, L., An Essay on the Nature and Significance of Economic Science, London: Darby Books, 1932.

NORMATIVE AND EMPIRICAL ISSUES IN ECONOMIC THEORY: A PHILOSOPHICAL EXAMINATION

by Douglas MacLean and Claudia Mills

Half a century ago, the logical positivists tried to separate sharply the normative from the empirical. Underlying positivism was a view about science which maintained a fundamental distinction between facts and values and which generated a correspondingly naive view about the value-free methods of science.

Most philosophers of science today would reject the idea that facts and values can be clearly distinguished. Values seep into science at many junctures, determining what to count as evidence, how to report findings, the bases for accepting and rejecting theories, and so on. Those who theorize about science are now more comfortable talking about the value of various scientific enterprises.

The relationship between facts and values and between normative and empirical claims, even in the physical sciences, is more complicated than the positivists believed. In the social sciences, especially where normative theories may have play an important role, these relationships raise crucial problems, as they do in philosophy.

Philosophers who construct normative theories may have a tendency to ignore relevant facts. They often make hand-waving appeals to their intuitions or to what "we believe," without explaining why these appeals should lend credence to the conclusions that follow. If we examine our normative convictions critically, perhaps even empirically, we can find many reasons not to be sanguine. It may be a regrettable fact that normative philosophy does not have better empirical grounding, but it is not surprising. Philosophers, after all, are not trained to work from empirical data. Often they think that theirs is an *a priori* discipline, like mathematics.

It is more remarkable, however, to notice similar tendencies in the work of social scientists. They too can be blinded by their normative convictions and led to make *a priori* claims about how people actually behave or about how they would behave under certain circumstances. We will demonstrate these tendencies below with examples.

Our interest in this subject has been stimulated by some recent work in what is now called behavioral decision theory. This is empirical work, much of which is designed explicitly to clarify the distinction between normative and empirical claims in economics, decision theory, and policy analysis. This work has general philosophical interest as well.

Douglas MacLean is Director and Claudia Mills is Associate of the Center for Philosophy and Public Policy, University of Maryland, College Park, MD.

Facts, Norms, and Intuitions

Normative philosophy and normative social science, that is, all those disciplines or sub-disciplines where moral, social, or political theory, or conceptions of rational thought and action play central roles, begin with intuitions. "Ethical theories have to start from somewhere," Bernard Williams writes, and after finding various other possible foundations for moral philosophy unsatisfactory, he concludes that "the only starting point left is ethical experience itself" (18).1/ We will argue to a similar conclusion.

By claiming that intuitions are the foundations of normative inquiry, however, we do not mean to suggest that these disciplines rely on some mysterious cognitive faculty that puts us directly in touch with true and objective values, as some earlier "intuitionist" moral theories may have suggested. We take intuitions to be, rather, as Williams describes them, our "spontaneous convictions, moderately reflective but not yet theorized" (18). Calling them spontaneous convictions suggests correctly that intuitions are usually invoked without being supported by appeals to evidence. Williams characterizes intuitive responses as "assured and unprompted" (18).

Our intuitions sometimes rely on empirical assumptions, however, about our psychological make-up, circumstances of our world, the range of politically feasible options, or our culture's shared moral experiences. They are to this extent vulnerable to empirical testing, and they are vulnerable in other ways as well. Intuitions may be the foundations of normative inquiry, but they are corrigible foundations. How are they corrigible? And how are empirical considerations relevant to their justification?

One obvious problem with intuitions is that our firmest convictions may not be as widely shared as we think; almost certainly they will not be as widespread as we think. Intuitions are culture-bound and shaped by our particular psychologies and historical circumstances. Philosophers' intuitions may be theirs alone, or they may be shared by few people beyond the circle of like-minded colleagues or those committed to the same theory.

To gather evidence is the only way to know whether normative convictions are in fact universally shared, or that they would be universal under certain ideal conditions. Gathering evidence is something philosophers are not prone to do; nor is it always easy to know where to look for evidence, even for those who are trained in empirical methods. How could we know, for example, whether people are happier if they are free? How can we determine how consumers would actually behave under ideal circumstances where there is perfect information and no transaction costs? Although our normative convictions about these matters clearly have empirical content, it is not at all clear what that content is.

Conflicting norms is another problem with intuitions. An investigation may be required to determine whether a conclusion supported by one set of intuitions might not be undermined by other equally plausible intuitions. But this will be an investigation into moral reasoning and justification, not an empirical survey. One philosophical critic of reliance on intuitions argues that "rival theories can be played off against intuitions about particular cases; and such intuitions can never be decisive" (6).

1/ Underscored numbers in parentheses refer to sources cited in References at the end of this article.

A third problem is that even our widely shared intuitions may have suspicious pedigrees. Peter Singer maintains that our firmest moral convictions derive from discarded religious systems, from warped views of sex and bodily functions, or from customs necessary for the survival of the group in social and economic circumstances that now lie in the distant past" (17). Relying on intuitions can make normative inquiry conservative and reactionary. According to this picture, we use our convictions to create theories to pat ourselves on the back for having the convictions we do.

Arguments that rest ultimately on intuitions might thus be circular or question-begging. Derek Parfit points out that if some theory "has been taught for more than two millennia, we must expect to find some echo in our intuitions...[It] cannot be justified simply by an appeal to the intuitions that its teaching may have produced" (13). Normative theories must be critical; they must allow us to abandon intuitions in the light of reason and to be prepared to accept counter-intuitive conclusions. As several decision theorists have pointed out, if our normative theory of choice was also an adequate descriptive account of our decision making behavior, it would serve no useful purpose (15).

Given these problems and tendencies, we might agree with the conclusion of yet another critic, that "it is puzzling why an intuition, a normative conviction, should be a test of anything" (3). Intuitions are clearly problematic and complicate the relations between empirical and normative claims, but we cannot simply reject them. We have no place else to begin. A further look at the nature of moral reasoning can give some indication for this.

Derek Parfit describes two ways of doing ethics or arguing about morality.

> One is the Low Road, that merely appeals to our intuitions. The other is the High Road, Meta-Ethics. If we can give the best account of the nature of moral reasoning, we can hope that this will imply particular claims about morality. We can hope that our Meta-Ethics will imply conclusions in Ethics (13).

Anti-intuitionist moral theorists, as Parfit suggests, repair to the high road of meta-ethics in the hope of developing a theory of moral reasoning that will allow them to rise above moral intuitions.

But we are not aware of any developed view about moral reasoning to convince us that the appeal to rationality alone will yield any substantive moral conclusions; and the history of moral philosophy, dating back at least to Kant, suggests that substantive moral claims cannot be derived from a theory of rationality or practical reasoning alone.

Accounts of moral reasoning, moreover, also appeal to intuitions; if not relatively concrete moral intuitions, then more abstract and consequently less well-grounded normative convictions about what is reasonable. In formal models of rational choice like expected utility theory, for instance, we are asked to accept as intuitive some rather abstract axioms about which most people have no firm convictions. Whatever normative appeal attaches to these axioms derives as much from the elegance and structure of the system of which they are a part as from our independent considered judgments about, for instance, the normative force of the substitution axiom or the sure thing principle. Granted, we are not likely to be illicitly importing intuitions that are vestiges from our premoral (or prerational) pasts into our acceptance of an axiomatic normative theory. But we are being invited, at least implicitly, to think that

mathematics provides the model for the structure of rational thought, and that rational decisions must involve maximizing something. If this is not an artifact of a discarded religion, it does at least show a cultural bias.

We do not mean to suggest that these theories must give us an incorrect model of moral reasoning, but only that they also appeal in the end to intuitions. There is no reason why we should place more confidence in abstract intuitions than in our more concrete, moral convictions. The latter will be firmer, even though they may also be tainted in ways the critics of intuition point out.

Finally, these more concrete moral norms also determine our understanding of rationality and our ideas about proper forms of moral reasoning. For example, whether reactive attitudes like regret and reproach are rational, or just how morality makes room for acting on purely personal rather than impersonal reasons, cannot be settled by considerations of moral reasoning alone. Our understanding about these important questions inevitably appeals to out moral intuitions.

Thus, we should look instead to ethical theories that do not try to avoid intuitions but attempt instead to incorporate them into deliberations in ways that respect their fallibility. John Rawls defends a theory of this kind (16). He describes moral reasoning as a process of trying to arrive at what he calls a "reflective equilibrium" between theory and intuition, which aims to save as many of our intuitions as possible while producing a rational structure of principles that makes clear which intuitions have to be rejected or modified. On this conception of morality, intuitions are corrigible, but they are also the basis upon which our principles and theories are constructed. "There is a definite, if limited, class of facts against which conjectured principles can be checked, namely our considered judgments in reflective equilibrium" (16).

It is a further feature of Rawls's method that in the process of developing a theory and arguing for principles that will save our intuitions and resolve conflicts, we must also bring in general facts about human psychology and political feasibility. Thus, empirical claims are explicitly incorporated into moral reasoning, and normative principles are tied to empirical facts.

We see no conflict between accepting intuitions as the starting point for normative philosophy and social science and accepting the criticisms we described earlier of a reliance on them. It means only that our method of reasoning must include procedures for criticizing and justifying intuitions that help us avoid the pitfalls that the critics describe.

We must recognize, first, that intuitions are empirically vulnerable. This is in part because of the way facts and norms are intermingled in reasoning about morality and rationality, but also because of the ways, as a matter of empirical fact, we come to have the normative convictions we do.

We must ensure that our intuitions are generally plausible. The best way to do this is to begin by relying on intuitions that are both widely believed and well supported by critical refection. If either of these conditions fails, then we are justified in suspending belief or in demanding further justification.

Finally, in order to avoid question-begging and circular reasoning, we must take a critical attitude toward the uses of intuitions in normative inquiry. Norms supply us with reasons, but these are reasons in the sense of

justification, not necessarily reasons in the sense of causes. Justification is a normative issue, about which we might have reasonable convictions; causes are empirical issues, not to be settled a priori.

Whatever force these prescriptions have is best appreciated by looking at some of the errors and confusions that occur when normative and empirical claims are not properly distinguished or are thoughtlessly combined in normative theories. Our main example will focus on an issue in economics and policy analysis, which is directly related to some interesting work in behavioral decision theory.

Theories on Social Cost

"The Problem of Social Cost" is a famous article by Ronald Coase that has altered the way many people think about law and economics (5). Coase was writing against the background of A.C. Pigou, whose well-known work in welfare economics in the 1920s led economic and policy theorists to become concerned with the "externalized" costs of economics activities (14). These costs, like pollution, are borne by society at large and lead to underpricing of commodities because they tend not to be counted as costs of production. Pigou regarded the social cost of production as the "actual" cost, to which he intended to include these externalities. He maintained that social costs should be internalized, or that prices should reflect actual costs, and he proposed that this could be accomplished only by legally imposed fines and bounties on externalities.

Pigou's suggestion for implementation was regarded as unfeasible, but legal reformers at the time did argue for internalizing at least those social costs that resulted in harm by replacing the traditional negligence standards in tort law with strict liability standards. Pigou and his followers thus maintained that producers should be held responsible for the full cost of production. This was an intuition or a normative conviction, which they attempted to justify on economic grounds.

Coase attacked this justification and thus Pigou's conception of social cost. He argued instead that the problem of social cost is a reciprocal one, which requires both a perpetrator and a victim (5).

> The question is commonly thought of as on in which A inflicts harm on B and what has to be decided is: how should we restrain A? But this is wrong...The real question that has to be decided is: should A be allowed to harm B or should B be allowed to harm A?

Now, this is clearly a normative question, and Coase addresses it by arguing for two conclusions. The first, which has come to be known as Coase's theorem, is that in an ideal situation, defined as a situation where there are no transaction costs, parties will bargain to the same result no matter how entitlements or liabilities are assigned. There "the damaging business is not liable for any of the damage which it causes," Coase writes, "the allocation of resources will be the same...as it was when the damaging business was liable for damage caused" (5).

In other words, Coase denies that the existence of externalities leads to underpricing. If A inflicts harm on B, then the harm will remain so long as the cost of the harm is less than the cost of removing it; if A is liable, then A will compensate B; and if A is not liable, B will accept the harm. If the cost of the harm exceeds the cost of removing it, then if A is liable, A will

remove the harm, and if A is not liable, B will pay to remove it. Income will be distributed differently, according to how the entitlements and liabilities are assigned, but no economic inefficiencies are involved. The prices of commodities will remain the same. Thus, Coase concludes, "The intimate result (which maximizes the value of production) is independent of the legal position if the pricing system is assumed to work without cost" (5).

This is an empirical, predictive claim. Coase's justification for it, however, consists not of economic data establishing that this is in fact how consumers or firms behave in ideal situations, but rather of a few simple hypothetical examples suggesting that this is how reasonable people ought to behave. His appeal, therefore, is to intuitions about reasonable economic behavior, intuitions that are clearly driven by a normative theory about rational choice and behavior. Coase does not acknowledge or attempt to defend this theory, perhaps (judging from his use of the indicative rather than the subjunctive mood) because he believes that actual behavior coincides with normative behavior, in other words, that the behavior of actual people under ideal market circumstances coincides with the behavior of ideal people under these same circumstances.

If this is Coase's reasoning, then it is formally acceptable. It relies reasonably on an intuition, granted certain assumptions. The coincidence of actual behavior or beliefs and relatively obvious norms, we have suggested, provides as firm a foundation for a normative argument as we can imagine. Coase obviously believed the norm is obvious and the corresponding behavior and beliefs actual.

That Coase relies on hypothetical examples rather than empirical data is not necessarily a weakness in his argument, for, as his second argument makes clear, his theorem has no practical application. There are virtually always transaction costs in the world in which we actually live. For the pricing system to work smoothly, as Coase puts it, the market must be regulated at least to the extent that rights and liabilities are clearly assigned. Often they are not, and it is always costly to make these corrections through regulation. His second argument, then, is that we cannot simply assume that the social cost of regulating to correct these defects will be less than the cost of not regulating. "What has to be decided," he concludes, "is whether the gain from preventing harm is greater than the loss which would be suffered elsewhere as a result of stopping the action which produces the harm" (5).

Although this appears to be a straightforward normative conclusion, it is qualified by other remarks Coase makes. He acknowledges that there may be moral reasons for assigning entitlements one way rather than another and that these may be worth some degree of inefficiency. He thus concurs with Frank Knight that "welfare economics must ultimately dissolve into a study of aesthetics and morals" (5). Moral considerations and the economic inefficiencies that result from acting on them are treated by Coase as separate, extraneous matters, not central to his principal argument: that rights and entitlements can be treated economically as factors of production whose distribution will not affect economic efficiency. Then, together with the assumption that the problem of social cost is reciprocal and thus no more the responsibility of the perpetrator than of the victim, economic activity should be regulated primarily to correct economic inefficiencies. Regulations should not be undertaken unless their expected benefits exceed their expected costs.

Our interest in Coase's argument lies in showing how his empirical assumptions about how people will behave under ideal circumstances is crucial to his normative convictions about the reciprocity of social costs and the justification of regulation. One of Coase's examples involves a cattle-raiser's herd damaging a farmer's crop. Coase maintains that whether the farmer is entitled to damages or not has no effect on the size of the herd the cattle-raiser will actually keep. "This will be the same as it would be if the cattle-raiser had to pay for damages caused by his cattle, since a receipt forgone of a given amount is the equivalent of a payment of the same amount" (5).

The claim that opportunity costs (receipts forgone) are equivalent to out-of-pocket costs (payments) is a cornerstone of the theory of rational choice to which Coase subscribes. This belief leads him to his assertion about how people do behave or how they will behave. But the empirical claim is false. And because it is false, Coase cannot appeal to it to justify the intuition that helps support his normative views. That the claim is false has been established by behavioral experiments, which we will discuss presently, but it has been known for some time (though perhaps not when Coase wrote his article) that opportunity costs are not empirically equivalent to out-of-pocket costs in the sense that people are not willing to sell a commodity or an entitlement for the same price as they would be willing to buy it. The equivalent is merely normative, and because it is, according to our suggestions for establishing the normative status and relevance of intuitions, it cannot simply be appealed to as an intuition but must have an independent justification.

Numerous empirical studies have demonstrated that buying and selling prices are not equivalent (12). In some attempts to measure indirectly the economic value of different kinds of environmental protection to hunters and other users of wilderness and wetlands, expressed selling prices are typically four to sixteen times higher than expressed willingness to pay (12). This effect has also been notices more directly in other contexts. For example, wage rate differentials for hazardous and nonhazardous occupations which are otherwise similar increased dramatically after 1970 when the Occupational Safety and Health (OSH) Act was passed (4). Presumably workers then regarded a safe workplace as their right and increased its value, as a selling price, over what they were willing earlier to pay for it.

Some legal theorists have been quick to embrace these findings and use them to argue against Coase. They have suggested rationalizations of the difference in buying and selling behavior that fit their moral intuitions. Duncan Kennedy writes that "people generally have greater concern for an attachment to things as they are than to things as they could be," (11) and that "people are disinclined to disrupt an equilibrium state" (11). He cites as "a matter of empirically demonstrable fact" our moral institution about the difference in strength between duties not to harm and duties to aid (11). It is not entirely clear in Kennedy's article how this intuition, even if it is widely held and reflected in behavior, explains the effect. Perhaps it can, but nowhere in the article or its many references is it indicated that this demonstrable fact has, in fact, been demonstrated. By assuming *a priori* that his intuition is reflected in consumer behavior, Kennedy commits the same mistake as Coase. Both are led by their normative convictions to make *a priori* claims about consumer behavior. With the benefit of hindsight, Kennedy correctly predicts the behavior, but that is no reason to think his intuitive account explains it correctly.

Mark Kelman, another legal theorist, similarly argues against Coase that legal entitlements will create and reinforce a sense of moral right, which will explain the relevant finding (10). He writes:

> If the consumer...values pure water more highly because he is entitled to it,...[h]e will not give up his right to pure water for the $100 cost of abatement because he values the right too highly. But pure water, a good like any other on the market that he can buy or not, may not be worth $100 to him. Perhaps society learns what to value in part through the legal system's description of our protected spheres.

Kelman makes two related claims here. The general claim is that a society's legal policies will have a causal impact on its citizens' evaluative sensibilities. The second assertion, an application of the more general claim, is that gaining an entitlement to a good will increase its perceived value.

The interesting general claim seems intuitively plausible. Of course we should look for confirming evidence, but it would be surprising if a social institution as important as a legal system did not have a powerful effect in shaping the moral consciousness of its citizens. Surely the image of blacks in the United States has changed as a result of civil rights legislation. But the truth of the application cited by Kelman is not so obvious that it can be taken a priori.

Kelman suggests that the consumer values the entitlement to pure water, which he is not willing to abandon lightly. But deciding to exchange pure water for $100 is to exercise an entitlement, not to abandon it. Why should the right increase the value of the water? Moreover, if the argument is a general one, it should be symmetrical: it should also be true that we value $100 more if we haye it (and the entitlement to it) than if we do not have it or are trying to get it. Perhaps we do, but it is not so obvious that we value the money we have more than the money we want. Or perhaps something about unspoiled rivers, wilderness, or other environmental goods can explain why their perceived value is affected by whether or not we believe we have entitlements to them. We will suggest later why this might be the case, but if true, this is the fact that needs to be explained.

These critics are mirror images of Coase. They commit the same mistakes. Normative convictions lead them both to make a priori claims about empirical matters, for example, the prediction or explanation of consumer behavior. Coase, it turns out, predicts incorrectly. We will now suggest that Kennedy and Kelman explain incorrectly as well.

The Explanation of Some Empirical Findings

Behavioral decision theory is a field that has been shaped by the work of psychologists Daniel Kahneman and Amos Tversky. Their behavioral theory, which they call prospect theory, attempts to demonstrate and explain some of the widespread violations of the axioms and assumptions of expected utility theory (7). Much of their work consists of showing how these violations are systematic.

Two features of prospect theory help to explain the behavioral effects involved in the difference between buying and selling. The first is the framing phenomenon. Kahneman and Tversky maintain that people do not attempt to maximize any kind of objectively defined utility function. Rather, our preferences tend to be determined by what we see as gains and losses, defined

in relation to some neutral reference point. This reference point may or may not be determined by something objective, like a person's current asset position. It may also be determined by other factors, such as the individual's expectations or by how a problem is presented or described.

Gains and losses, as measured from the reference point, are treated differently. The utility function for losses is steeper for most people than it is for gains. This feature or prospect theory is called loss aversion. It also turns out that people tend to be risk seeking in the domain of losses (people prefer a gamble to a sure loss with a similar "expected value") and risk averse in the domain of gains (people prefer a sure gain to a gamble).

Together, these features can explain certain behavior that fundamentally violates expected utility theory. If framing determines whether prospects are viewed as gains or losses, and if gains and losses are valued differently, then preferences can change with different decision frames. Kahneman and Tversky's well-known examples illustrate this effect (8, 9).

Asian Disease

> Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:
>
> Survival Frame
>
> If Program A is adopted, 200 people will be saved. [72%]
>
> If Program B is adopted, there is a 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved. [28%]
>
> Mortality Frame
>
> If Program C is adopted 400 people will die. [22%]
>
> If Program D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die. [78%]

When the prospects are framed in terms of survival, the majority of respondents prefer A to B, but describing the same prospects in terms of mortality induces the majority of people to take up a different reference point, from which the different treatment of gains and losses leads to a reversal of preferences.

The fact that people can be manipulated into making inconsistent responses by clever descriptions of a decision problem is not in itself exciting or important. We should be concerned, however, with the explanation of the cognitive processes that lead to these reversals. This is the crucial empirical issue.

Now let us consider two intuitively plausible explanations. The first, suggested by some philosophers, appeals to the moral intuitions mentioned by Kennedy. Most people, we assume, think the obligation to avoid killing is stronger than any obligation to save lives. We must not kill innocent people, even for great social gain; yet we are not required to accept great sacrifices in order to save lives that would otherwise be lost, were we to do nothing. Not everybody accepts these intuitions, but most people do, or so this

explanation goes. Thus, when the Asian Disease example is presented in the survival frame, people are inclined to think of 200 lives that would be saved if Program A is chosen but that would be put at significant risk by choosing Program B. If these 200 lives are lost, they are lost as a result of our decision to choose B over A, and the moral prohibition against causing 200 deaths is more stringent than the obligation to try to prevent the deaths of the additional 400. In the choice between C and D, however, people are inclined to focus on the 600 deaths that will be caused in the absence of any combative program. Since we have, it is assumed, only a relatively weak obligation to prevent deaths caused through no fault of our own, it is less important to ensure that 200 are saved and more attractive to take the gamble of saving the full 600. After all, whatever action we take here is to our moral credit.

Having suggested this explanation of these different preferences by appealing to moral intuitions, few philosophers, we think, would go on to defend this particular reversal. After all, when the two frames are presented simultaneously, the alternatives can be seen to involve identical choices under different descriptions. Moral intuitions are appealed to, therefore, in order to give a causal or behavioral explanation of the common responses. The intuitions may be normatively justifiable; the behavior they give rise to in this instance is not.

Moral intuitions doubtless play a significant role in determining preferences and choices. The stronger the normative appeal of the intuition, one might think, the stronger this causal role will be. This is precisely where it is important to distinguish normative from empirical issues. The causal role of moral intuitions is an empirical issue; as such, it needs to be tested. As we shall see, the tests made by Kahneman, Tversky, and other experimenters following their lead suggest that this is not the best explanation of the preference reversal in the Asian Disease example.

A different explanation of loss aversion comes from decision theorists attempting to reconcile the normative theory with empirical results. Several of these revisionists have suggested that the concept of utility has to be broadened to include not just the utility of the outcomes of choice, in this instance, the utility of money or of lives saved, but also of the "cognitive" factors that attach to the act of choosing (2). Specifically, these theorists have suggested that people are strongly influenced by the potential for regretting their choices when risks turn out badly. People will give up some expected value of outcomes in order to avoid potential regret. Regret avoidance may be a strong human motivation. It may even be rational to avoid regret or blame, whether or not the regret or blame itself has a rational foundation. But it is seldom made explicit in these reconciliation attempts whether regret avoidance is part of a normative theory of rational choice or part of an empirical theory that explains behavior.

Regret avoidance is an intuitively plausible way to explain some important empirical findings (such as the well-known "Allais paradox"). But regret avoidance and the moral explanation both turn out to be ad hoc. Because both have independent normative appeal, it is easy to assume that people act according to them and thus that they apply causally as well. To see the problem with this reasoning, consider another of Kahneman and Tversky's examples.

Conditional Gambles

Assume yourself richer by $300 than you are today. You are offered a choice between

a sure gain of $100 [72%]
50% chance to gain $200, and 50% chance to gain nothing [28%]

Assume yourself richer by $500 than you are today. Your are forced to choose between

a sure loss of $100 [36%]
50% chance to lose nothing and 50% chance to lose $200 [64%]

The moral explanation is rendered implausible by the Conditional Gambles example. The same kind of reversal that occurs in Asian Disease also occurs when the stakes are money rather than lives and when the consequences affect the welfare of the actor alone. Nor can potential regret explain this set of preferences. Regret theories cannot distinguish the two problems in Conditional Gambles, because they assume the prospects are evaluated in relation to each other, rather than to some common reference point. In Conditional Gambles, the difference between the consequence of one's choice and the consequence of having chosen differently is identical in both problems. Prospect theory gives a more basic and general explanation of the empirical findings than the moral explanation or the appeal to potential regret.

Further evidence from a number of other, very simple experiments seems to reinforce this conclusion. In one experiment, for example, Knetsch and Sinden gave half their subjects tickets to a lottery and the other half $3. When the first group was given an opportunity to sell their tickets for $3, 82 percent kept them, but when the second group was allowed to buy lottery tickets for their $3, only 38 percent wanted the tickets (12). Simple experiments like this one undermine the relatively sophisticated or morally appealing explanations of preference reversals or buying and selling differentials in more complicated contexts. Appeals to the value of entitlements, rights, or potential regret have an intuitive plausibility in the richer contexts, but little plausibility here. What sensitivity to the value of the entitlements in question could conceivably explain why people tend to keep whichever of the lottery ticket or the money they are initially given? Loss aversion is a more convincing explanation for these simple cases, and it is sufficient to explain the more interesting and important cases as well. It is the simplest and best explanation of some very general behavior.

We do not mean to suggest that other factors cannot also influence our decisions, but parsimony and other theoretical considerations should lead us to the more basic explanation, which can apply to a wider range of phenomena. The evidence suggests that the best empirical account of the difference in buying and selling behavior is cognitively simple: a combination of framing effects and loss aversion. If moral principles, the value of legal entitlements, or potential regret also influence our behavior in these cases, as perhaps we think they should, then different and further empirical evidence will be needed to establish this fact.

Normative Considerations

We have been concerned so far with making two basic points. The first is about the importance of distinguishing normative from empirical claims. We have noted some errors and fallacies that occur when these are conflated. In particular, it is a mistake to think that because some claim has considerable normative appeal, it must be part of the explanation of human behavior. A

normative justification is a reason for thinking that some principle ought to help determine a person's preferences and decisions, but it is not a reason for thinking that it actually does. Normative claims are justified by moral reasoning, which appeals to intuitions. Causal claims are empirical claims, and they need to be tested and justified by appeal to evidence and data. People do not always behave according to their own normative convictions; otherwise phenomena like weakness of the will could not occur. The confusions that result when normative and behavioral claims are not properly distinguished can also have important implications for policies, as they have in discussions of the problem of social cost.

Our second point has been to argue that prospect theory gives the best explanation of a set of important behavioral violations of normative theories. Framing effects, combined with loss aversion, are sufficient to explain some of the complicated and important cases, and this explanation is confirmed by some simple experiments in which more subtle hypotheses, for example, appealing to moral intuitions, income effects, the consciousness-raising effects of entitlements, and so forth, plainly fail to apply. Generality is a virtue of an explanation. If a single hypothesis can explain behavior as complicated as the difference in wage rates for hazardous work before and after the passage of the OSH Act, as well as effects as simple as the fact that if you give someone a $2 pen, she probably will not sell it for its market value, but if you give her $2, she probably would not buy the pen, then this generality tends to confirm the hypothesis.

This leaves us with an important question: What should we say normatively about loss aversion? Recall that in our discussion of the problem of social cost, some morally plausible intuitions would have led to the same behavior predicted by loss aversion. Our argument there concerned the best empirical explanation of the difference between buying and selling prices. But we might appeal to our normative convictions to justify that behavior. Thus, loss aversion might in some cases be morally defensible or even praiseworthy. And if behavior is morally defensible, then it is also rational, if by rational we simply mean reasonable and not something more technical or narrow like the strict pursuit of self-interest or maximizing expected utility according to the axioms of a normative theory of choice.

At the same time, loss aversion is a far too widespread and general phenomenon simply to embrace as a moral or rational norm. We respond with the majority in the Asian Disease example, but we confess that this preference shift cannot be justified. It shows only how most people can be manipulated by the framing of a decision problem. We would hope to be able to overcome this kind of manipulation. A clever experimenter could probably lead us to keep a lottery ticket, or to refuse to buy one, in the same way that most people can be so manipulated, but we find no normative justification for the loss aversion demonstrated in these cases. So framing and loss aversion are to be neither simply defended nor attacked on normative grounds. Loss aversion appears to be a basic psychophysical reaction that may be reasonable in some contexts but unreasonable in others. The moral issues will not be settled by this behavior. They must be treated independently, and we must appeal to other normative convictions in doing so.

Where should this moral inquiry begin? Our discussion suggests that it might well start with issues involved in framing and entitlements. These are closely related, for as the empirical work in behavioral decision theory reminds us, entitlements will affect preferences and are thus necessary to determine what

count as efficient outcomes. The conditions of efficiency depend on the assumptions of entitlements that go into the framing of the problem.

Some will no doubt think that, as a normative matter, this should not make any difference. Buying and selling prices should be equivalent. We should, as Coase recommends, be able to take and "opportunity cost approach" to policy questions and "compare the total product yielded by alternative social arrangements"; and we should be able to determine opportunity costs independently from entitlements. Generally speaking, according to this way of thinking, the different ways of framing outcomes, probabilities, or acts ought normatively to be equivalent.

We disagree. "A finger twitch that initiates a sequence of events resulting in a human death" and "a murder" are not morally equivalent descriptions of an event, even though in some sense both descriptions may be true of the same event. In morality and in the taw, murder and causing a death are quite different. Issues about responsibility and punishment often turn on which of competing descriptions, equivalent in some sense, is appropriate. Context, intentions, and other moral factors, therefore, can determine how a problem ought to be framed. Independent normative considerations will determine which framings or descriptions are morally or rationally equivalent.

Entitlements will be essential factors in these considerations. To take a dramatic example, consider the morally unproblematic right not to be raped. The importance and value of this right cannot possibly be reduced or made equivalent to what we might determine that citizens in some given social circumstances ought rationally to pay for protection against rape. It would be morally outrageous to suggest that the latter measure can be used to determine full compensation for rape victims or that the incidence of rape should be the same in societies, regardless of whether citizens paid to protect victims from rape or whether rapists were allowed to rape so long as they were willing to compensate their victims. A society that allocated moral and legal entitlements in this latter way would be acting according to a moral code that we would find incomprehensible.

This example suggests how we might understand the consciousness-raising feature of entitlements. We found it puzzling why a legal entitlement to pure water should increase the value we place on the water, as Kelman suggested; we could not see why the value of something to a person ought to be affected by her owning it. It would be normatively unreasonable and empirically bizarre to think that owning something is necessary to appreciate its value fully, that we perhaps undervalue whatever we do not own. When Cecil Rhodes wrote, "Expansion is everything...I would annex the planets if I could," (1) we understand this to be not a sensitive desire to appreciate other cultures and people more highly, but simply the ranting of a mad imperialist. The suggestion is also counterintuitive. It is the grass we do not own that is commonly supposed always to be greener.

The consciousness-raising effect of entitlements comes when we stop thinking of all legal entitlements and moral rights on the model of property rights, which we fully and properly exercise when we buy and sell. The right not to be raped is not merely the right a person has to her own body as her property. This is why violations of that right cannot be regarded as a legal issue of takings, which is justified if compensation is offered. The right not to be raped, like many other rights, derives not from ownership but from the dignity of the person. It does not arise from a legal convention, as the right to drive on the right-hand side of the street does. Rather, the laws of any civilized

society must protect that right because of its independent moral importance. The question of how much we ought to spend as a society to offer what degree of protection against rape is a separate and difficult policy question that will be informed by the moral severity of rape. The amount we determine is appropriate cannot be regarded as full market value for the unconsented use of a woman's body.

Likewise, coming to think we are entitled to live in a land with areas of wilderness protected for posterity, or to a workplace that is safe from certain health risks, might have similar consciousness-raising effects. We might come to see these as moral rights of a kind that cannot strictly be regarded on the model of property rights. We might see ourselves as having special duties to conserve our environmental legacy for future generations; we might rethink the respect we are owned as members of the nation's work force; and so on. This will lead to an appreciation of rights that might well and reasonably affect our willingness to exchange these things for other benefits. If so, then it will be highly misleading simply to attach some financial increment to a good as an expression of the value we place on our entitlement to it. Reducing all rights to property tights will not allow us to conduct the proper kind of normative investigations into these very important areas of social policy.

A defensible theory of moral rights, which did not reduce all rights to property rights, would also have to demonstrate how the protection of rights does not call for too great a sacrifice of welfare, and it would probably have to explain how rights have compelling but not absolute force and when they can be overridden because of other considerations. We have not tried to present such a full-blown conception of moral rights here. But we have tried to show, in a general way, how moral issues must be treated independently of patterns of economic behavior and how certain moral rights and entitlements might influence our other preferences and decisions.

References

1. Arendt, H., The Origins of Totalitarianism. New York: Meridian, 1958, p. 124.

2. Bell, D., "Regret in Decision Making Under Uncertainty," Operations Research 30 (1982), pp. 961-81.

3. Brandt, R., A Theory of the Good and the Right. Oxford: Clarendon Press, 1979, p. 21.

4. Brown, C., "Equalizing Differences in the Labor Market," Quarterly Journal of Economics 94 (1980), pp. 113-34.

5. Coase, R., "The Problem of Social Cost," Journal of Law & Economics 3 (1960), pp. 1-44.

6. Kagan, S., "The Present-Aim Theory," Ethics 96 (1986), p. 757.

7. Kahneman, D. and Tversky, A., "Prospect Theory: An Analysis of Decision Under Risk," Econometrica 47 (1979), pp. 263-91.

8. "The Framing of Decisions and the Psychology of Choice," Science 211 (January 30, 1981), pp. 453-58.

9. _____. "Rational Choice and the Framing of Decisions," Journal of Business 59 (1986), pp. 251-78.

10. Kelman, M., "Consumption Theory, Production Theory, and Ideology in the Coase Theorem," Southern California Law Review 52 (1979), p. 695.

11. Kennedy, D., "Cost-Benefit Analysis and Entitlement Problems: A Critique," Stanford Law Review 33 (1981), pp. 401-02.

12. Knetsch, J. and Sinden, J.A., "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value," Quarterly Journal of Economics 98 (1984), pp. 507-21.

13. Parfit, D., Reasons and Persons. Oxford: Clarendon Press, 1984, p. 130.

14. Pigou, A. C., The Economics of Welfare. London: Macmillan, 1920.

15. Raiffa, H., Decision Analysis. Reading, Mass.: Addison Wesley, 1968, p. x.

16. Rawls, J., A Theory of Justice. Cambridge, Mass.: Harvard University Press, 1971.

17. Singer, P., "Sidgwick and Reflective Equilibrium," The Monist 58 (1974), pp. 490-517.

18. Williams, B., Ethics and the Limits of Philosophy. Cambridge, Mass.: Harvard University Press, 1985, pp. 93-97.

POSTSCRIPT: THE NEW SPIRIT OF INQUIRY IN ECONOMICS AND PHILOSOPHY

Michael McPherson

Economists may have suffered at times from too great a willingness to take philosophers' word for how they should do their business. Economists' varied attempts over the last forty years to live up to the canons of "logical positivism" surely had some of that quality(1).1/ Today the climate is rather different. Philosophy no longer seems so certain of its own ground, and economists, while by no means comfortable about the state of their own discipline, are skeptical that any methodological recipes for advance exist.

This mutual caution has, perhaps surprisingly, opened the way for a new and potentially very fruitful type of collaboration between economists and philosophers. One way of putting it is that a number of philosophers are at the moment quite willing to "worry with" the economists about problems of mutual interest, while a number of economists are eager to learn from this kind of joint inquiry, an inquiry very different in spirit from any kind of "preaching" by philosophers to economists. The philosophers' interest is stimulated in part by recognition that neither the philosophy of knowledge nor that of morals can be a purely a priori discipline. Both must be informed by the "facts" in the broadest sense about how people gather knowledge and how they develop moral judgments. Since economists work systematically at acquiring knowledge and improving moral judgment about one important aspect of human affairs, philosophers interested in these matters must "get their hands dirty" by learning close up how these cognitive and moral enterprises work. Economists are acutely aware, more than they have been for some time, of the crudity and limits of their available tools and of the fragility of their claims to knowledge. Anyone willing to think with them about the very hard problems the discipline must deal with, and help them better understand the limits and the possibilities of the economists' perspective, is encouraged to participate.

This spirit of mutual curiosity, of engagement in a process of inquiry where no one has settled answers, was greatly in evidence during the sessions Gene Wunderlich organized (two of which I was able to attend) and in the papers growing out those sessions that are published here. It is a spirit that Dan Hausman and I have seen in abundance since we founded the Cambridge University Press journal, Economics and Philosophy, three years ago.

As the papers in this collection make clear, collaborations between economists and philosophers offer no golden road to enlightenment. Philosophy helps, when it helps, by clearing ground, raising new questions, and providing new perspectives on the structure of problems. One way of helping is by destroying unwarranted certainties, as in Buchanan's critique of careless uses of the

---

Michael McPherson is a Professor, Department of Economics, Williams College, Williamstown, MA.

1/ Underscored numbers in parentheses refer to sources cited in the Reference at the end of this article.

among different aspects of thought, as in MacLean and Mills' sorting out the normative and empirical dimensions of critiques of rational decision making. Joe Pitt illustrates another service philosophers can provide, that of unfolding the complexities buried within a seemingly simple idea; in Pitt's case the conception of the consumer as economically "rational."

It is an exciting time to be engaged with economics and philosophy. Although reading these papers may remind us of how little we know and how hard economics is, they can also remind us of how much we have to learn, and how many promising paths are open.

Reference

1. Hausman, Dan, "Economic Methodology and Philosophy of Science," The Boundaries of Economics. Ed. Gordon C. Winston, Cambridge, England: Cambridge University Press, forthcoming.

*U.S. GOVERNMENT PRINTING OFFICE:1987-180-916:40740/ERS